A weight function theory of positive order basis function interpolants and smoothers

Phillip Y. Williams

5/31 Moulden Court, Canberra, ACT 2617, Australia.

philwill@pcug.org.au, phil.williams@canberra.edu.au

February 5, 2008

I would like to thank Dr Dave Davies and the University of Canberra for their generous assistance. Also, thanks to my Masters degree supervisors Dr Markus Hegland and Dr Steve Roberts of the Centre for Mathematics and its Applications (CMA) at the Australian National University.

ABSTRACT

In this document I develop a weight function theory of positive order basis function interpolants and smoothers.

In Chapter 1 the basis functions and data spaces are defined directly using weight functions. The data spaces are used to formulate the variational problems which define the interpolants and smoothers discussed in later chapters. The theory is illustrated using some standard examples of radial basis functions and a class of weight functions I will call the tensor product extended B-splines.

Chapter 2 shows how to prove functions are basis functions without using the awkward space of test functions So,n which are infinitely smooth functions of rapid decrease with several zero-valued derivatives at the origin. Worked examples include several classes of well-known radial basis functions.

The goal of Chapter 3 is to derive 'modified' inverse-Fourier transform formulas for the basis functions and the data functions and to use these formulas to obtain bounds for the rates of increase of these functions and their derivatives near infinity.

In Chapter 4 we prove the existence and uniqueness of a solution to the minimal seminorm interpolation problem. We then derive orders for the pointwise convergence of the interpolant to its data function as the density of the data increases.

In Chapter 5 a well-known non-parametric variational smoothing problem will be studied with special interest in the order of pointwise convergence of the smoother to its data function. This smoothing problem is the minimal norm interpolation problem stabilized by a smoothing coefficient.

In Chapter 6 a non-parametric, scalable, variational smoothing problem will be studied, again with special interest in its order of pointwise convergence to its data function.

Contents

0.1 Introduction

In this document I develop a weight function theory of **positive order** basis function interpolants and smoothers. Note that the Appendix of this document contains a list of basic notation, definitions and properties also used in this document.

This document had its genesis in the development of a *scalable algorithm* for *Data Mining* applications. Data Mining is the extraction of complex information from large databases, often having tens of millions of records. Scalability means that the time of execution is linearly dependent on the number of records processed and this is necessary for the algorithms to have practical execution times. One approach is to develop additive regression models and these require the approximation of large numbers of data points by surfaces. Here one is concerned with approximating data by surfaces of the form $y = f(x)$, where $x \in \mathbb{R}^d$, $y \in \mathbb{R}$ and d is any dimension. Smoothing algorithms are one way of approximating surfaces and in particular we have decided to use a class of non-parametric smoothers called basis function smoothers, which solve a variational smoothing problem over a semi-Hilbert space of continuous *data functions* and express the solution in terms of a single *basis function.* Indeed, several years ago I succeeded in developing a scalable smoothing algorithm(unpublished) of the basis function type. This algorithm was derived by approximating a minimal smoothing problem on a regular rectangular grid. In this document we develop some theoretical tools to construct and analyze it. This theoretical approach applies in any dimension but the smoothing algorithm is only practical up to three dimensions. In higher dimensions the matrices are too large to put into computer memory.

I started my Masters degree (supervised by Dr. Markus Hegland and Dr. Steve Roberts at the ANU, Canberra, Australia) searching for a scalable basis function smoothing algorithm and had the good fortune to devise such an algorithm by approximating, on a regular grid, the convolution in Definition 38 of the space J_G in Dyn's review article [4]. For the case of zero order basis functions this algorithm is analyzed in Chapter 4 of the document Williams [22], and the positive order basis function case is studied in Chapter 6 of this document.

Dr. Hegland was particularly interested in using the tensor product hat(triangular) function as a basis function. At about this time we had a visit by the late Professor Will Light who showed me his paper [11] which defined basis functions in terms of weight functions using the Fourier transform. Light and Wayne's weight function properties, given in Definition 3 below, were designed for positive order basis functions which excluded the tensor product hat function. They were designed for the well-known 'classical' radial weight functions. I therefore developed a version of his theory designed to generate zero order basis functions, including both tensor product and radial types. This theory was developed in Chapter 1 of Williams [22] and requires that the basis functions have Fourier transforms which can take zero values outside the origin since this is a property of hat functions.

Chapter by chapter:

Chapter 1 The goal of this chapter is to extend the theoretical work of Light and Wayne in [11] to allow classes of weight functions defined using integrals and which allow tensor product weight functions. Two *weight function* definitions are introduced, one designed for the standard radial weight functions and one designed for tensor product functions. These definitions allow for basis functions which have Fourier transforms which are zero on a closed set of measure zero and are analogous to the zero order definitions introduced in Williams [22]. These definitions involve the positive integer order parameter θ and the smoothness parameter κ which can take any non-negative real value.

A class of tensor product weight functions are introduced which I call the *extended B-splines.*

Both the semi-inner product *data space* $\left(X_w^\theta, |\cdot|_{w,\theta}\right)$ and the *basis functions* G are defined in terms of the weight function. We then prove some completeness and smoothness properties of the data space as well as some smoothness and positive definiteness properties of the basis function. In fact the data space consists of continuous functions and the basis functions are continuous, and the data space is used to define the variational interpolation and smoothing problems of later chapters and the basis functions are used to express their (unique) solutions.

The *extended B-spline* basis functions are shown to be the convolutions of hat (triangle) functions.

Chapter 2 shows how to prove functions are basis functions without using the awkward test functions $S_{\emptyset,2\theta} = \{\phi \in S : D^\alpha \phi(0) = 0,\ |\alpha| < 2\theta\}$ where S is the C^∞ space of rapidly decreasing functions given in Definition 243. Worked examples include several classes of well-known radial basis functions, the choice here following Dyn [4]: the thin-plate splines, the shifted thin-plate splines, the multiquadric and inverse multiquadric functions and the Gaussian. In the last section I will illustrate the method using a non-radial example: the fundamental solutions of homogeneous elliptic differential operators.

The goal of Chapter 3 is to derive 'modified' inverse-Fourier transform formulas for basis functions and the data functions and to use these formulas to obtain bounds for the rates of increase near infinity of these functions and their derivatives. This will be done by proving a general inverse-Fourier transform formula for a subspace of the distributions and then applying it to both the basis functions and the data functions. Some other basis function properties are derived from their weight function properties.

Chapter 4 studies the *minimum seminorm interpolation problem.* Topics include unisolvency and Lagrange polynomial interpolation and related operators and matrices, the existence and uniqueness of the basis function solution, a matrix equation for the solution and the pointwise convergence of the interpolant to its data function on a bounded region.

We use a *data function* $f_d \in X_w^\theta$ and *unisolvent independent data* $X = \left\{x^{(k)}\right\}_{k=1}^N$ contained in a bounded *data region* Ω to generate N data points $\left(x^{(k)}, f_d\left(x^{(k)}\right)\right)$ and the interpolation problem requires us to minimize $|\cdot|_{w,\theta}$ over all $f \in X_w^\theta$ which interpolate these data points. The seminorm $|\cdot|_{w,\theta}$ is converted into a *Light norm* $\|\cdot\|_{w,\theta}$ and the problem is reformulated as the *minimum norm interpolation problem.* Geometric Hilbert space theory using orthogonal projections shows the solution, denoted $\mathcal{I}_X f_d$, is unique. Indeed, this solution is identical to that of the seminorm problem and it lies in the finite dimensional basis function space $W_{G,X}$ (Definition 130). It is called a *basis function solution.*

We are now interested in what happens to the interpolant as the data X becomes denser in Ω when we use the *cavity measure* of data density

$$h_X = \sup_{x \in \Omega} \operatorname{dist}(x, X).$$

Theorem 152 establishes an error estimate of order $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$ for the interpolant in the sense that there are positive constants k_I and h such that

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d|_{w,\theta}\, k_I\, (h_X)^\eta, \quad x \in \overline{\Omega},\ f_d \in X_w^\theta, \tag{1}$$

when $h_X < h$.

Note that in the case of the thin-plate, the shifted thin-plate and the Gaussian radial basis functions the Light and Wayne weight function definition of [11] yields the same orders of convergence as obtained above. However, slightly improved orders of convergence are obtained later: see Corollary 157 and examples where an improvement of 1/2 is obtained for the shifted thin-plate splines.

Chapter 5 studies the well-known *Exact smoother problem* (our terminology) which stabilizes the interpolant by adding a smoothing coefficient to the seminorm functional. Topics include the existence and uniqueness of the basis function solution i.e. the *Exact smoother*, a matrix equation for the solution and the convergence of the solution to its data function.

Using the same data as the interpolation problem the Exact smoother problem requires us to minimize the functional

$$J_e[f] = \rho\, |f|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^N \left| f\left(x^{(i)}\right) - f_d\left(x^{(i)}\right)\right|^2, \tag{2}$$

for $f \in X_w^\theta$. The proof will be carried out within a Hilbert space framework by formulating (Definition 168) the smoothing functional in terms of an inner product $(u,v)_V = \rho \langle u_1, v_1 \rangle_{w,\theta} + \frac{1}{N}(\widetilde{u}_2, \widetilde{v}_2)_{\mathbb{C}^N}$ on the product space $V = X_w^\theta \otimes \mathbb{C}^N$ and the continuous operator $\mathcal{L}_X : X_w^\theta \to V$ defined by $\mathcal{L}_X f = \left(f, \widetilde{\mathcal{E}}_X f\right)$ where $\widetilde{\mathcal{E}}_X f = \left(f\left(x^{(k)}\right)\right)$. We can now write $J_e[f] = \left\|\mathcal{L}_X f - \left(0, \widetilde{\mathcal{E}}_X f\right)\right\|_V^2$ and using the technique of orthogonal projection it is shown in Theorem 171 that a unique solution to the smoothing problem exists given by $s_e = \frac{1}{N}\left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f_d$. Like the interpolant, the Exact smoother lies in the finite dimensional basis function space $W_{G,X}$ and a matrix equation 5.20 is derived for the coefficients of the basis functions.

To derive our estimates for the Exact smoother error $|s_e(x) - f_d(x)|$ on a bounded data region we make use of the operators $\mathcal{L}_X$ and $\mathcal{L}_X^*$. Now $\|\mathcal{L}_X f\|_V$ is an equivalent norm to $\|f\|_{w,\theta}$ which implies that X_w^θ is also a reproducing kernel Hilbert space under the inner product $(\mathcal{L}_X f, \mathcal{L}_X f)_V$. It then follows that there exists a unique $R_{V,x} \in V$ such that $f(x) = (\mathcal{L}_X f, R_{V,x})_V$ and $R_x = \mathcal{L}_X^* R_{V,x}$. Thus, using the

geometric properties of the orthogonal projection

$$\begin{aligned}|s_e(x)-f_d(x)| = \left|(\mathcal{L}_X(s_e-f_d), R_{V,x})_V\right| &\le \|\mathcal{L}_X(s_e-f_d)\|_V \|R_{V,x}\|_V \\ &\le \left\|\mathcal{L}_X f_d - \left(0, \widetilde{\mathcal{E}}_X f_d\right)\right\|_V \|R_{V,x}\|_V \\ &= |f_d|_{w,\theta}\sqrt{\rho}\,\|R_{V,x}\|_V .\end{aligned}$$

We then estimate $\|R_{V,x}\|_V$ and show (eq'n 5.45) that there exist positive constants k_e, k_e' and h' such that

$$|s_e(x)-f_d(x)| \le |f_d|_{w,\theta}\, k_e\left(k_e'(h_X)^\eta + \sqrt{N_X\rho}\right), \quad x\in\overline{\Omega},\ f_d\in X_w^\theta,$$

when $h_X < h'$. Here X has N_X points and $k_e k_e' = k_I$. The latter equation means that when $\rho = 0$ the smoother estimate coincides with the interpolant error estimate. As with the interpolant estimates, slightly higher orders of convergence can be obtained - see Corollary 157 and examples.

Chapter 6 studies the *scalable Approximate smoother* (our suggested name). Topics include the derivation of the *Approximate smoother problem* and the existence, uniqueness of its solution i.e. the Approximate smoother, a matrix equation for the smoother, its scalability and the pointwise convergence of the smoother to the Exact smoother and to its data function.

Using data generated by a data function $f_d\in X_w^\theta$ and a **fixed** set $X'\subset\mathbb{R}^d$ the Approximate smoother problem requires us to minimize the Exact smoother functional 2 over the finite dimensional space $W_{G,X'}$. Using Hilbert space techniques and then matrix techniques the smoother s_a will be shown to exist, be unique and satisfy the regular matrix equation 6.10 i.e. if $s_a(x) = \sum\limits_{i=1}^{N'}\alpha_i' G(x-x_i') + \sum\limits_{i=1}^{M}\beta_i' p_i(x)$ then

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'} & G_{X',X}P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix}\begin{pmatrix}\alpha' \\ \beta' \\ \gamma'\end{pmatrix} = \begin{pmatrix} G_{X',X} \\ P_X^T \\ O_{M,N}\end{pmatrix} y, \tag{3}$$

where: $G_{X',X'} = G\left(x_i'-x_j'\right)$, $G_{X',X} = G\left(x_i'-x^{(j)}\right)$, $G_{X',X} = G\left(x^{(i)}-x_j'\right)$, $P_X = \left(p_j\left(x^{(i)}\right)\right)$, $P_{X'} = \left(p_j\left(x_i'\right)\right)$ and $\{p_i\}$ is a basis for the polynomials P_θ of degree less than θ.

The *Approximate smoother matrix is* $N+N'$ square and this dependence on the number of data points N implies *computational scalability* i.e. the computational effort required to construct and solve the matrix equation depends linearly on the number of data points N.

For a bounded data region Ω we first derive some uniform, pointwise convergence results which do not involve orders of convergence e.g. in Corollary 229 it is shown that as the points in X' get closer to those in X the Approximate smoother converges uniformly to the Exact smoother on $\overline{\Omega}$.

Orders of pointwise convergence are derived for the convergence of the Approximate smoother to the Exact smoother e.g. estimate 6.39 says that for some constants $B, C, r > 0$

$$|s_e(x)-s_a(x)| \le |f_d|_{w,\theta}\left(1 + B\frac{(h_{X'})^{\eta_G}}{\sqrt{\rho}}\right)\left(B(h_X)^{\eta_G} + C\sqrt{N_X\rho}\right), \quad x\in\overline{\Omega},\ f_d\in X_w^\theta, \tag{4}$$

when $h_{X'}, h_X < r$. Here $h_X = \sup\limits_{\omega\in\Omega}\operatorname{dist}(\omega, X)$ and $h_{X'} = \sup\limits_{\omega\in\Omega}\operatorname{dist}(\omega, X')$ measure the density of the point sets X and X'.

These estimates are then added to the Exact smoother error formulas of Chapter 5 to obtain error estimates for the Approximate smoother e.g. estimate 6.49.

Notation 1 *Regarding notation for distributions, in [11] Light and Wayne used an operator notation which I feel significantly reduced the readability of an already difficult paper, especially replacing* $(ix)^\gamma$ *by* $V_\gamma(x)$ *and replacing the exponential* e^{ixy} *by* $e_x(y)$. *For me this turned old mathematical friends into strangers. So I have settled on the notion of an* ***action variable*** *e.g. in the equations*

$$[\xi^\alpha f, \phi] = [f, \xi^\alpha \phi], \quad \left[e^{i\xi} f, \phi\right] = \left[f, e^{i\xi}\phi\right].$$

involving a distribution f *and a test function* ϕ, ξ *will be called the action variable. I will also use the dot notation e.g.*

$$\left[|\cdot|^2 f, \phi\right] = \left[f, |\cdot|^2\phi\right], \quad [f(\cdot - x), \phi] = [f, \phi(\cdot + x)].$$

My aim is to retain as much as possible of ***the form*** *of 'conventional' mathematical expressions.*

1

Extensions of Light's class of weight functions

1.1 Introduction

The goal of this chapter is to extend the theoretical work of Light and Wayne in [11] to allow classes of weight functions analogous to the zero order weight functions developed in Chapter 1 of Williams [22] i.e. weight functions which can generate positive order tensor product basis functions and the radial basis functions which have Fourier transforms with zeros outside the origin. A class of weight functions which I call the extended B-splines of positive order is used to illustrate the weight function properties and the basis function theory. Besides the basis function theory I have shown Light and Wayne's semi-Hilbert function spaces are still valid for the expanded weight function class but I give a different proof of completeness that uses single mappings in a manner analogous to Sobolev space theory.

The material of this chapter can be summarized as:

1. Define the positive order weight function properties and present equivalent properties and relationships between them. Introduce the *extended B-spline weight functions* which are tensor product functions with bounded support.

2. Define the semi-inner product space of continuous *data functions* and prove its completeness and smoothness properties.

3. Define the basis distributions of positive order and prove continuity and positive definiteness properties. Derive convolution formulas for the extended B-spline basis functions.

The theory of this document lays the foundations for the study of the basis function interpolation and smoothing problems introduced in later chapters.

1.2 An extended class of positive order weight functions

In this section we introduce our extended class of positive order weight functions.

1.2.1 Zero order weight functions and basis functions

In the Chapter 1 [22] Williams developed a theory of zero order weight functions and basis functions. Here we shall summarize the relevant parts of this theory which only needs simple L^1 Fourier theory e.g. Section 2.2 of Petersen [16]. However, here we will allow non-integer values of the smoothness parameter κ to match the positive order weight function definitions below.

Definition 2 ***Zero order weight functions with smoothness parameter*** κ

A weight function's properties are defined with reference to a set $\mathcal{A} \subset \mathbb{R}^d$ *which is closed and has measure zero. A weight function* w *is a mapping* $w : \mathbb{R}^d \to \mathbb{R}$ *which has the properties:*

W1 *There exists a closed set $\mathcal{A}$ with measure zero such that w is continuous and positive outside $\mathcal{A}$ i.e. $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ and $w > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$.*

W2 *For some (possibly non-integer) number $\kappa \geq 0$*

$$\int \frac{|x|^{2s}}{w(x)} dx < \infty, \quad 0 \leq s \leq \kappa. \tag{1.1}$$

The *basis function of order zero* is defined by

$$G = \left(\frac{1}{w}\right)^{\vee}. \tag{1.2}$$

It follows directly from Corollary 2.12 of Petersen [16] that $G \in C_B^{(\lfloor 2\kappa \rfloor)}$ and so I sometimes call κ the weight function *smoothness parameter.*

1.2.2 Motivation for the extended weight function class properties

I will now state the Light and Wayne weight function properties extracted from Section 3 of Light and Wayne [11]. To this list **I have added property B3.5 - not defined by Light and Wayne**. The space X of functions mentioned in this definition were used to define the variational interpolant. The basis function is used to construct the interpolant.

Definition 3 ***Light and Wayne's weight function properties*** *[11]*
A ***weight function*** *w is a mapping $w : \mathbb{R}^d \to \mathbb{R}$ which has properties A3.1 and A3.2 below, as well as combinations of the other properties. Properties A3.1 and A3.2 are used to define the semi-inner product space X of distributions.*

A3.1 $w \in C^{(0)}(\mathbb{R}^d \backslash 0)$.

A3.2 $w(x) > 0$ *on* $\mathbb{R}^d \backslash 0$.

Properties A3.3 and A3.4 are used to prove the completeness of the X space and to allow the definition of a basis distribution.

A3.3 $1/w \in L^1_{loc}$.

A3.4 *There exist μ, $R > 0$ and $C_R > 0$ such that*

$$\frac{1}{w(x)} \leq C_R |x|^{-2\mu}, \quad |x| \geq R.$$

B3.5 *We say that w has property B3.5 for parameters μ and* ***order*** *θ if w satisfies property A3.4 for some μ satisfying $\mu + \theta > d/2$.*

I have constructed Property B3.5 *from Light and Wayne [11] and it allows the space X to be embedded in the continuous functions $C^{(k)}$ where k is the largest integer such that $k < \mu + \theta - d/2$ (Theorem 2.18), and the basis distributions of order θ to be continuous functions $C^{(j)}$ where j is the largest integer such that $j < 2(\mu + \theta) - d$ (Theorem 3.14). Property B3.5 allows these weight function properties to be more easily compared with the 'extended' properties given below.*

Light and Wayne's weight function properties were designed to generate the positive order radial basis functions and the aim of this section is to extend Light and Wayne's weight function properties so that they can generate positive order tensor product basis functions and also to formulate the weight function properties in terms of integrals.

In Chapter 1 Williams [22] a class of weight functions, which I called the extended B-spline weight functions, was used to illustrate the weight function properties and the basis function theory. In this document a class of weight functions which I will call the *extended B-spline weight functions of positive order* is used to illustrate the weight function properties and the basis function theory.

In more detail, the extended weight function properties given below in Definition 4 were formulated by taking the following considerations into account:

1 I was interested in properties expressed in terms of integrals. Also, these properties should be scalable.

2 When a function w_1 defined on $\mathbb{R}^1$ has the weight function properties every tensor product $\prod\limits_{i=1}^{d} w_1(x_i)$ should also have these properties but perhaps for different parameter values. Now suppose w is a tensor product function of the 1-dimensional function w_1 i.e.

$$w(x) = \prod_{i=1}^{d} w_1(x_i), \quad x \in \mathbb{R}^d,$$

and suppose that w_1 has Light and Wayne properties A3.1 and A3.2. Now we want w to be a weight function i.e. satisfy A3.1 and A3.2. In general this is not possible since if w_1 is discontinuous at zero then w is discontinuous on the closed set $\bigcup\limits_{i=1}^{d} \{x : x_i = 0\}$. This motivates extended weight function property W1 of Definition 4 i.e. there exists a closed set $\mathcal{A} \subset \mathbb{R}^d$ of measure zero such that $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ and $w > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$.

3 I want to generalize the zero order extended B-spline weight functions studied in Chapter 1 Williams [22] to the positive order case. This will mean allowing basis functions with Fourier transforms which have zeros outside the origin i.e. weight function with poles (discontinuities) outside the origin. Indeed, the one-dimensional hat function is defined by

$$\Lambda(x) = \begin{cases} 1 - |x|, & |x| \leq 1, \\ 0, & |x| > 1, \end{cases} \qquad x \in \mathbb{R}^1, \tag{1.3}$$

and in higher dimensions it is defined as the tensor product

$$\Lambda(x) = \prod_{i=1}^{d} \Lambda(x_i), \quad x \in \mathbb{R}^d. \tag{1.4}$$

It is well known that

$$\widehat{\Lambda}(\xi) = \prod_{i=1}^{d} \widehat{\Lambda}(\xi_i) = (2\pi)^{-d/2} \prod_{i=1}^{d} \left(\frac{\sin(\xi_i/2)}{\xi_i/2}\right)^2, \quad \xi \in \mathbb{R}^d. \tag{1.5}$$

The zero order extended B-spline weight functions were created by generalizing the right side of 1.5 to the two-parameter class of functions

$$\frac{1}{w(\xi)} = \prod_{i=1}^{d} \frac{\sin^{2n} \xi_i}{\xi_i^{2k}}, \quad n, k \geq 1, \tag{1.6}$$

and then restricting the choice of parameters l and n. The positive order weight functions will involve a different choice of these parameters. Here $1/w$ has zeros outside the origin on a closed set of measure zero.

4 Property A3.4 is $1/w \in L^1_{loc}$ and so is already defined by an integral and this becomes extended weight function property W2.1 of Definition 4. Property A4.4 states: there exists $\mu \in \mathbb{R}^1$ and $R, C_R > 0$ such that $\frac{1}{w(x)} \leq C_R |x|^{-2\mu}$ for $|x| \geq R$. Now Lemma 50 requires that $1/w \in S'$ i.e. $1/w$ be a tempered distribution (Appendix A.5) and from part 2 of Appendix A.5.1 we see that A3.4 and A3.4 imply that $1/w$ is a regular, tempered distribution. Further it is clear that extended property W2.2 i.e. $\int\limits_{|\cdot| \geq r_2} \frac{1}{w|\cdot|^{2\sigma}} < \infty$ for some $\sigma > 0$ and $r_2 > 0$, combined with property W2.1 also implies that $1/w \in S'$. Happily, property W2.2 is a good substitute for property A3.4.

5 Some clarification of the relationship between property B3.5 and the new properties W3.1, W3.1* and W3.2 will be provided by Theorem 6 in Subsection 1.2.4.

1.2.3 The extended class of positive order weight functions

The weight function class of Light and Wayne will be extended as follows:

Definition 4 ***The weight function properties***

The properties are defined with reference to a set $\mathcal{A} \subset \mathbb{R}^d$ which is a closed set of measure zero. The weight function is a mapping $w : \mathbb{R}^d \to \mathbb{R}$ which has at least property W1:

W1 $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ *and* $w > 0$ *on* $\mathbb{R}^d \backslash \mathcal{A}$.

Property W1 will be used to define the semi-inner product distribution space X_w^θ for positive integer ***order*** *θ. This property is identical to the zero order weight function property W1 (Definition 2 above).*

W2 *Property W2 is satisfied if the following two sub-properties are satisfied:*

W2.1 $1/w \in L_{loc}^1$.

W2.2 $\int\limits_{|\cdot| \geq r_2} \frac{1}{w|\cdot|^{2\sigma}} < \infty$ *for some* $\sigma > 0$ *and* $r_2 > 0$.

Property W2 is used to prove completeness and C^∞ density results regarding the X_w^θ spaces, as well as allowing the definition of the basis distributions. Condition W2 is satisfied iff the function λ defined by:

$$\lambda(x) = \begin{cases} 0, & |x| \leq r_2, \\ \sigma, & r_2 < |x|, \end{cases} \tag{1.7}$$

satisfies $\int \dfrac{dx}{w(x)\, |x|^{2\lambda(x)}} < \infty$. *The function λ is useful for the concise expression of integral inequalities involving weight functions e.g. Lemma 26.*

Now to introduce properties W3.1, W3.1 and W3.2. These will allow the X_w^θ spaces to be continuously embedded in the continuous functions and the basis distributions to be continuous functions. We call W3.1* the* ***strong version*** *of W3.1. Suppose $\theta \geq 1$ is a positive integer and $\kappa \geq 0$:*

W3.1 *w has property W3.1 for order θ and parameter κ if there exists a multi-index α such that $|\alpha| = \theta$ and*

$$\int \frac{|x|^{2t}\, x^{2\beta}}{w(x)\, x^{2\alpha}} dx < \infty, \quad \begin{cases} |\beta| \leq \lfloor \kappa \rfloor, \\ 0 \leq t \leq \kappa - \lfloor \kappa \rfloor. \end{cases} \tag{1.8}$$

W3.1* *w has property W3.1* for order θ and κ if it has property W3.1 for θ and κ and* ***all*** *α such that $|\alpha| = \theta$.*

W3.2 *w has property W3.2 for order θ and κ if there exists $r_3 > 0$ such that*

$$\int\limits_{|x| \geq r_3} \frac{|x|^{2t}}{w(x)\, |x|^{2\theta}} dx < \infty, \quad 0 \leq t \leq \kappa.$$

W3 *If w has property W3.1 or W3.1* or W3.2.*

Remark 5

1. *This definition permits the weight function to have* ***discontinuities outside the origin*** *on a closed set $\mathcal{A}$ of measure zero. This allows the Fourier transform of the* ***basis function to have zeros outside the origin****. For example, the extended B-spline tensor product weight functions introduced in Subsection 1.2.8 will generate basis functions which have zeros outside the origin on a set of measure zero.*

2. *The extended properties have been framed with a desire to use* ***integrals to formulate properties*** *which generalize Light and Wayne's properties.*

3. *Like the Light and Wayne properties A3.3 and A3.4, the extended properties W2.1 and W2.2 ensure that $1/w$* ***is a regular tempered distribution*** *(or generalized function of slow growth) as per part 2 of Appendix A.5.1.*

4. *Properties W3.1 and W3.1* are designed for use with* ***tensor product weight functions*** *such as the tensor product extended B-spline weight functions introduced in Subsection 1.2.8, and the equivalent property of part3 Theorem 7 may be even more suitable e.g. Theorem 13.*

   With regard to point 2 of Subsection 1.2.2 above, when a weight function w_1 defined on $\mathbb{R}^1$ has any of the weight function properties W2.1, W2.2, W3.1 every tensor product $\prod_{i=1}^{d} w_1(x_i)$ also has the same properties but perhaps for different parameter values.

5. *Property W3.2 is clearly designed to* ***handle radial weight functions****.*

6. *In one dimension properties W3.1 and W3.1* are identical but in higher dimensions property W3.1* may be stronger that property W3.1. However, property W3.1* is easier to handle analytically but W3.1 will give better results.*

7. *The expressions used to define properties W3.1 and W3.2 may look a little strange because we have not combined the* ***exponents from the denominator and the numerator*** *but these expressions allow direct comparison of the two definitions and point to important similarities between them. Also, if a weight function has property W3 for order θ then we will only define the function spaces X_w^θ and the basis functions for this order.*

8. *I turns out that $X_w^\theta \subset C^{\lfloor\kappa\rfloor}$ and $G \in C^{\lfloor 2\kappa\rfloor}$ so we shall sometimes refer to κ as the* ***smoothness parameter****.*

9. *Observe that all the* ***weight function properties are scalable****, w.r.t. both the dependent and independent variables.*

10. *If we set $\theta = 0$ in the definition of the extended properties we obtain a set of* ***zero order weight function*** *properties: property W1 for the positive order weight functions is identical to the zero order weight function property W1 of Definition 2. Properties W3.1 and W3.1* become equivalent (by Theorem 7) to zero order property W2. Properties W3.2 and W2 are equivalent to zero order property W2. When $\theta = 0$ positive order property W2 stays the same and ensures that $1/w$ is a tempered distribution. It allows the definition of a* ***zero order basis distribution*** *which corresponds to the positive order distributions of Definition 44.*

1.2.4 Relationships between the extended properties and Light and Wayne's weight function properties

The next theorem justifies the use of the terminology *extended weight function properties.*

Theorem 6 *Suppose for parameters $\mu \in \mathbb{R}^1$ and $R > 0$ a weight function w satisfies properties A3.1 to A3.4 of Light and Wayne's Definition 3. Then:*

1. *w satisfies the extended weight function properties W1 and W2 of Definition 4 with $\mathcal{A} = \{0\}$, when $\sigma > d - 2\mu$, $\sigma \geq 0$ and $r_2 = R$.*

2. *Suppose w also satisfies property B3.5 for μ and order θ i.e. $\mu + \theta > d/2$. Then the weight function satisfies property W3.2 of Definition 4 for order θ and all κ such that $0 \leq \kappa < \mu + \theta - d/2$.*

Proof. Part 1 Clearly $\mathcal{A} = \{0\}$ is a closed set of measure zero. It is also clear that w has properties W1 and W2. Now to prove property W4. Property A3.4 requires that for some $R > 0$

$$\frac{1}{w(x)} \leq C_R |x|^{-2\mu}, \quad \text{when } |x| \geq R.$$

Regarding the definition of property W4, choose $\sigma > d - 2\mu$, $\sigma \geq 0$ and $r = R$. Then

$$\int\limits_{|x|\geq r} \frac{dx}{|x|^\sigma w(x)} \leq C_R \int\limits_{|x|\geq r} \frac{dx}{|x|^{\sigma+2\mu}} < \infty,$$

since $\sigma + 2\mu > d$. Thus w has property W4.

Part 2 Choose any $r' > 0$. Then, since w has property A3.4

$$\int_{|x|\geq r'} \frac{|x|^{2\kappa}\,dx}{|x|^{2\theta}\,w(x)} = \int_{|x|\geq r'} \frac{dx}{|x|^{2(\theta-\kappa)}\,w(x)} \leq C_R \int_{|x|\geq r'} \frac{|x|^{-2\mu}\,dx}{|x|^{2(\theta-\kappa)}} = C_R \int_{|x|\geq r'} \frac{dx}{|x|^{2(\mu+\theta-\kappa)}}.$$

But the condition $0 \leq \kappa < \mu + \theta - d/2$ implies $\mu + \theta - \kappa > d/2$ and so the last integral exists. ■

1.2.5 Equivalent criteria for properties W3.1, W3.1* and W3.2

Theorem 7 *The following three criteria are equivalent to weight function property W3.1:*

1. $\int \frac{|x|^{2\lambda}}{x^{2\alpha}w(x)}dx < \infty, \quad 0 \leq \lambda \leq \kappa.$
2. $\int \frac{|x|^{2\lambda}}{x^{2\alpha}w(x)}dx < \infty, \quad \lambda = 0,\ \kappa.$
3. $\int \frac{x_i^{2t}}{x^{2\alpha}w(x)}dx < \infty, \quad t = 0,\ \kappa$ *and* $i = 1, \ldots, d.$

Proof. Part 1 Use the identity $|x|^{2\theta} = \sum\limits_{|\beta|=\theta} \frac{\theta!}{\beta!} x^{2\beta}$.

Part 2 A simple proof is obtained by splitting the domain of integration about $|x| = 1$.

Part 3 Since $s \geq 0$ implies there exist constants a_s, $b_s > 0$, independent of x, such that

$$a_s \sum_{i=1}^{d} x_i^{2s} \leq |x|^{2s} \leq b_s \sum_{i=1}^{d} x_i^{2s}, \quad x \in \mathbb{R}^d,$$

this part is clearly equivalent to part 2. ■

Remark 8 *The previous theorem clearly also applies to weight function property W3.1* where we assume the inequalities are true for all α satisfying $|\alpha| = \theta$.*

Theorem 9 *The following criteria are equivalent to weight function property W3.2 for order $\theta \geq 1$and $\kappa \geq 0$:*

1. $\int\limits_{|x|\geq r_3} \frac{|x|^{2(\kappa-\lfloor\kappa\rfloor)}\, x^{2\beta}}{|x|^{2\theta}\, w(x)} dx < \infty, \quad |\beta| = \lfloor\kappa\rfloor.$
2. $\int\limits_{|x|\geq r_3} \frac{x_k^{2(\kappa-\lfloor\kappa\rfloor)} x^{2\beta}}{|x|^{2\theta}\, w(x)} dx < \infty, \quad \begin{cases} |\beta| = \lfloor\kappa\rfloor, \\ k = 1, \ldots, d. \end{cases}$

Proof. Part 1 is proved by using the identity

$$|x|^{2\kappa} = |x|^{2(\kappa-\lfloor\kappa\rfloor)}\, |x|^{2\lfloor\kappa\rfloor} = |x|^{2(\kappa-\lfloor\kappa\rfloor)} \sum_{|\beta|=\lfloor\kappa\rfloor} \frac{\lfloor\kappa\rfloor!}{\beta!} x^{2\beta} = \sum_{|\beta|=\lfloor\kappa\rfloor} \frac{\lfloor\kappa\rfloor!}{\beta!} |x|^{2(\kappa-\lfloor\kappa\rfloor)}\, x^{2\beta},$$

and part 2 is proved by applying the inequalities,
$\frac{1}{d} \sum\limits_{k=1}^{d} x_k^{2p} \leq |x|^{2p} \leq \sum\limits_{k=1}^{d} x_k^{2p}, \quad 0 \leq p \leq 1$, with $p = \kappa - \lfloor\kappa\rfloor$ to the criterion of part 1. ■

1.2.6 Relationships between the extended weight function properties

The next result demonstrates some relationships between the weight function property W2 and properties W3.1 and W3.2.

Theorem 10 *Suppose w has weight function property W1. Then:*

1. *Property W3.1 implies property W3.2 for any $r_3 > 0$.*

2. *Property W3.1 implies property W2 for any $\sigma \geq \theta$ and any $r_2 > 0$.*

3. *Property W3.2 implies property W2.2 for $\sigma = \theta$ and any $r_2 \geq r_3$.*

4. *Property W3.1* implies property W3.1.*

Proof. Instead of property W3.1 we use the equivalent criterion of part 1 of Theorem 7 i.e. there exists a multi-index α satisfying $|\alpha| = \theta$ such that $\int \frac{|x|^{2t} dx}{w(x) x^{2\alpha}} < \infty$ for all $0 \leq t \leq \kappa$.

Part 1 Choosing any $r_3 > 0$

$$\int\limits_{|x| \geq r_3} \frac{|x|^{2t} dx}{w(x) |x|^{2\theta}} = \int\limits_{|x| \geq r_3} \frac{|x|^{2t}}{w(x) \sum\limits_{|\mu|=\theta} \frac{\theta!}{\mu!} x^{2\mu}} dx \leq \int\limits_{|x| \geq r_3} \frac{|x|^{2t}}{w(x) \frac{\theta!}{\alpha!} x^{2\alpha}} dx = \frac{\alpha!}{\theta!} \int\limits_{|x| \geq r_3} \frac{|x|^{2t}}{w(x) x^{2\alpha}} dx < \infty,$$

by property W3.1.

Part 2 We prove properties W2.1 and W2.2. Suppose K is compact. Then for some $r > 0$, $K \subset B(0; r)$ and

$$\int\limits_{|x| \leq r} \frac{dx}{w(x)} = \int\limits_{|x| \leq r} \frac{x^{2\alpha} dx}{w(x) x^{2\alpha}} \leq \int\limits_{|x| \leq r} \frac{|x|^{2\theta} dx}{w(x) x^{2\alpha}} \leq r^{2\theta} \int\limits_{|x| \leq r} \frac{dx}{w(x) x^{2\alpha}}.$$

The last integral exists since w has property W3.1. Thus $\frac{1}{w} \in L^1_{loc}$.

We next show that W2.2 is true for $\sigma = 2\theta$ and any $r_2 > 0$. In fact, the identity $|x|^{2\theta} = \sum\limits_{|\beta|=\theta} \frac{\theta!}{\beta!} x^{2\beta}$ implies $|x|^{2\theta} \geq \frac{\theta!}{a!} x^{2\alpha}$, and so

$$\int\limits_{|x| \geq r_2} \frac{dx}{|x|^{2\theta} w(x)} \leq \frac{\alpha!}{\theta!} \int\limits_{|x| \geq r_2} \frac{dx}{x^{2\alpha} w(x)}.$$

The last integral then exists by property W3.1 with $t = 0$ and $\beta = 0$.

Part 3 By inspection of the integrals defining property W3.2.

Part 4 Directly from the definition of W3.1* in Definition 4. ■

1.2.7 *Other general results*

The next result gives an integral which exists for all weight functions with property W3.

Theorem 11 *Suppose a weight function w has property W3 for order θ and parameter κ. Then*

$$\int\limits_{|x| \geq r_4} \frac{|x|^{2s} dx}{w(x) |x|^{2\theta}} < \infty, \quad 0 \leq s \leq \kappa, \tag{1.9}$$

where $r_4 = 0$ if w has property W3.1 and $r_4 = r_3$ if w has property W3.2.

Proof. $\boxed{\textbf{Case 1}\text{: } w \text{ has property W3.1}}$ Then $r_4 = 0$ and by part 1 Theorem 7 there exists a multi-index α such that $|\alpha| = \theta$ and $\int\limits_{|x| \geq r_4} \frac{|x|^{2s} dx}{w(x) x^{2\alpha}} < \infty$ when $0 \leq s \leq \kappa$. But the identity $|x|^{2\theta} = \sum\limits_{|\alpha|=\theta} \frac{k!}{\alpha!} x^{2\alpha}$ of part 5 Appendix 242 implies $\frac{k!}{\alpha!} x^{2\alpha} < |x|^{2\theta}$ when $x \neq 0$ so that $\frac{\alpha!}{k!} \int\limits_{|x| \geq r_4} \frac{|x|^{2|\gamma|}}{w(x) x^{2\alpha}} > \int\limits_{|x| \geq r_4} \frac{|x|^{2|\gamma|} dx}{w(x) |x|^{2\theta}}$ and the integral 1.9 exists.

$\boxed{\textbf{Case 2}\text{: } w \text{ has property W3.2}}$ Then $r_4 = r_3$ and $\int\limits_{|x| \geq r_4} \frac{|x|^{2\kappa} dx}{w(x) |x|^{2\theta}} < \infty$. But if $0 \leq s \leq \kappa$ then $\int\limits_{|x| \geq r_4} \frac{|x|^{2s} dx}{w(x) |x|^{2\theta}} = \int\limits_{|x| \geq r_4} \frac{1}{|x|^{2(\kappa - s)}} \frac{|x|^{2\kappa} dx}{w(x) |x|^{2\theta}} \leq \frac{1}{{r_4}^{2(\kappa - s)}} \int\limits_{|x| \geq r_4} \frac{|x|^{2\kappa} dx}{w(x) |x|^{2\theta}} < \infty$ so the integral 1.9 exists. ■

1.2.8 The extended B-spline weight functions

We will now introduce the extended B-spline weight functions. These are two-parameter, tensor product weight functions and in the next theorem we will derive necessary and sufficient conditions for a weight function to have property W3.1*.

Theorem 12 *Suppose $d \geq 1$, $\kappa \geq 0$ is real and $\theta \geq 1$ is an integer. For given integers $l, n \geq 0$ the tensor product, extended B-spline weight function w is defined by*

$$w(x) = \prod_{i=1}^{d} \frac{x_i^{2n}}{\sin^{2l} x_i}. \tag{1.10}$$

Then when $d = 1$ there exists a closed set $\mathcal{A}_d$ of measure zero such that w is a function with properties W1 and W2. It also has property W3.1 for order θ and κ iff n and l satisfy*

$$n \geq \kappa + 1 - \theta, \qquad l \geq n + \theta. \tag{1.11}$$

When $d > 1$ there exists a closed set $\mathcal{A}_d$ of measure zero such that w is a function with properties W1 and W2. It also has property W3.1 for θ and κ iff n and l satisfy*

$$n \geq \kappa + 1, \qquad l \geq n + \theta. \tag{1.12}$$

Proof. Clearly for each parameter pair l, n there exists a closed set $\mathcal{A}_d$ of measure zero such that property W1 is satisfied. Further, since property W3.1* implies W3.1, by parts 1 and 2 of Theorem 10 w has property W2 and so we need only prove property W3.1* i.e. for all α such that $|\alpha| = \theta$

$$\int \frac{x_i^{2t}}{x^{2\alpha} w(x)} dx < \infty, \quad t = 0, \kappa; \; i = 1, \ldots, d.$$

Now

$$\int \frac{x_i^{2t}}{x^{2\alpha} w_s(x)} dx = \int \frac{x_i^{2t}}{x^{2\alpha}} \prod_{i=1}^{d} \frac{\sin^{2l} x_i}{x_i^{2n}} dx = \prod_{i=1}^{d} \int \frac{s^{2t}}{s^{2\alpha_i}} \frac{\sin^{2l} s}{s^{2n}} ds, \tag{1.13}$$

and these integrals will exist
iff they exist near the origin and near infinity,
iff $2t + 2l - 2\alpha_i - 2n > -1$ and $2n + 2\alpha_i - 2t > 1$ for $t = 0, \kappa$ and all i,
iff $2l - 2\alpha_i - 2n > -1$ and $2n + 2\alpha_i - 2\kappa > 1$ for all i,
iff $\kappa - n + 1/2 < \alpha_i < l - n + 1/2$ for all i,
iff $\kappa - n + 1/2 < \alpha_i \leq l - n$ for all i.

If $d = 1$ then $\alpha_i = \theta$ and so w has property W3.1* iff

$$\kappa - n + 1/2 < \theta \leq l - n. \tag{1.14}$$

If $d > 1$ then the integrals 1.13 exist whenever $|\alpha| = \theta$,
iff $\kappa - n + 1/2 < \alpha_i \leq l - n$ for all i, whenever $|\alpha| = \theta$, and
$\kappa - n + 1/2 < 0 \leq l - n$ and $\kappa - n + 1/2 < \theta \leq l - n$,
iff $\kappa - n + 1/2 < \alpha_i \leq l - n$ for all i, whenever $|\alpha| = \theta$, and
$\kappa - n + 1/2 < \alpha_i \leq l - n$ for $\alpha_i = 0, \theta$,
iff $\kappa - n + 1/2 < \alpha_i \leq l - n$ for all i, whenever $|\alpha| = \theta$, and
$\kappa - n + 1/2 < 0$ and $\theta \leq l - n$,
iff

$$\kappa - n + 1/2 < 0 \text{ and } \theta \leq l - n, \tag{1.15}$$

an 1.14 and 1.15 imply 1.11 and 1.12 of the theorem. ■

We now show that better results may be achieved using property W3.1 i.e. for given parameters θ and κ a larger combination of values of n and l satisfy property W3.1 than W3.1*. However when $\theta < d$ the results are the same.

Theorem 13 *Suppose w is the extended B-spline weight function 1.10 with parameters l and n.*

Then w is a weight function with property W2. It also has property W3.1 for given order $\theta \geq 1$ and $\kappa \geq 0$ iff n and l satisfy

$$n > \kappa + 1/2 - \min \alpha, \qquad l \geq n + \max \alpha, \tag{1.16}$$

for some a such that $|\alpha| = \theta$.
Further, w *satisfies property W3.1 for order* θ *and* κ *iff*

$$n > \kappa + 1/2 - \lfloor \theta/d \rfloor, \qquad l \geq n + 1 + \lfloor \theta/d \rfloor, \tag{1.17}$$

and this corresponds to any of the permutations of the $\alpha = (\alpha_i)$ *defined by 1.18 of the proof.*

Proof. For this theorem we use the criterion for property W3.1 of part 3 Theorem 7 i.e. there exists a multi-index α such that $|\alpha| = \theta$ and $\int \dfrac{x_i^{2t} dx}{x^{2\alpha} w(x)} < \infty$ for $t = 0, \kappa$ and $i = 1, \ldots, d$.

Suppose that $|\alpha| = \theta$. Then all the integrals $\int_{\mathbb{R}^1} \frac{s^{2t} \sin^{2l} s}{s^{2\alpha_i} s^{2n}} ds$ must exist for $t = 0, \kappa$,
iff integral exists near zero and near infinity,
iff $2\alpha_i + 2n - 2t - 2l < 1$ and $2\alpha_i + 2n - 2t > 1$ for $t = 0, \kappa$,
iff $\alpha_i + n - t - l < 1/2$ and $\alpha_i + n - t > 1/2$ for $t = 0, \kappa$,
iff $\alpha_i + n - l < 1/2$ and $\alpha_i + n > 1/2$ and $\alpha_i + n - \kappa - l < 1/2$ and $\alpha_i + n - \kappa > 1/2$,
iff $1/2 - n < \alpha_i < l - n + 1/2$ and $1/2 - n + \kappa < \alpha_i < l - n + \kappa + 1/2$,
iff $1/2 - n + \kappa < \min \alpha$ and $\max \alpha < l - n + 1/2$,
iff $1 - n + \kappa \leq \min \alpha$ and $\max \alpha \leq l - n$,
iff $n \geq 1 + \kappa - \min \alpha$ and $l \geq n + \max \alpha$,
which proves 1.16.

Next we want to show that n and l satisfy property W3.1 iff they satisfy 1.17 i.e. iff they satisfy 1.16 for α given by

$$\alpha_i = \begin{cases} 1 + \lfloor \theta/d \rfloor, & 1 \leq i \leq \operatorname{rem}(\theta, d), \\ \lfloor \theta/d \rfloor, & \operatorname{rem}(\theta, d) < i \leq d, \end{cases} \tag{1.18}$$

or any of its permutations. Clearly if n and l satisfy 1.17 they satisfy 1.16 for α given by 1.18. The converse will be true if we can show that the α given by 1.18 simultaneously solve the problems $\max\limits_{|\alpha|=\theta} \min \alpha$ and $\min\limits_{|\alpha|=\theta} \max \alpha$ because this would mean that if α satisfies 1.17 then it must satisfy 1.16. If $\lfloor \theta/d \rfloor = \theta/d$ then these problems are clearly solved by 1.18 i.e. by $\alpha_i = \theta/d$. Now suppose $\lfloor \theta/d \rfloor < \theta/d$ i.e. $\operatorname{rem}(\theta, d) > 0$. If $\min \alpha \geq 1 + \lfloor \theta/d \rfloor$ then $|\alpha| \geq d + d \lfloor \theta/d \rfloor = d + \theta - \operatorname{rem}(\theta, d) > \theta$, and if $\max \alpha \leq \lfloor \theta/d \rfloor$ then $|\alpha| < d \lfloor \theta/d \rfloor = \theta - \operatorname{rem}(\theta, d) < \theta$. Thus $\min \alpha \leq \lfloor \theta/d \rfloor$ and $\max \alpha \geq 1 + \lfloor \theta/d \rfloor$ and so 1.18 solves both $\max\limits_{|\alpha|=\theta} \min \alpha$ and $\min\limits_{|\alpha|=\theta} \max \alpha$. ■

In Section 1.6.4 we will derive expressions for the basis functions of positive order generated by the weight functions of the last theorem.

1.3 The spaces $S_{\emptyset,n}$ related spaces

We now define some key spaces and exhibit some of their properties.

Definition 14 ***The spaces*** $S_{\emptyset,n}$ ***and*** $C_{\emptyset,n}^{\infty}$

$$S_{\emptyset,0} = S, \qquad S_{\emptyset,n} = \{\phi \in S : D^{\alpha}\phi(0) = 0, \ |\alpha| < n\}, \ n = 1, 2, 3, \ldots, \tag{1.19}$$

and we endow $S_{\emptyset,n}$ *with the subspace topology induced by the space* S. S *is the space of* C^{∞} *functions of rapid decrease used as test functions for the tempered distributions.* S *is endowed with the countable seminorm topology described in Appendix 243.*

$$C_{\emptyset,0}^{\infty} = C^{\infty}, \qquad C_{\emptyset,n}^{\infty} = \left\{\phi \in C^{\infty} : D^{\beta}\phi(0) = 0, \ |\beta| < n\right\}, \ n = 1, 2, 3, \ldots,$$

so the space $C_{\emptyset,n}^{\infty}$ *retains the constraints of* $S_{\emptyset,n}$ *near the origin.*

The next result gives a simple upper bound for functions in $S_{\emptyset,n}$ near the origin. This follows directly from the estimate A.5 (Appendix A.8) of the integral remainder term of the Taylor series expansion.

$$|u(x)| \leq \left(\sum_{|\alpha|=n} \|D^{\alpha}u\|_{\infty} \right) |x|^{n}, \quad u \in S_{\emptyset,n}, \ x \in \mathbb{R}^{d}. \tag{1.20}$$

Noting that C_{BP}^{∞} is the space of C^{∞} functions for which each derivative is bounded by a polynomial we will need the following useful results:

Theorem 15

1. *If $f, g \in C_{BP}^{\infty}$ and $\psi \in S$ then $f\psi \in S$ and $fg \in C_{BP}^{\infty}$.*

2. *$S_{\emptyset,k} = S \cap C_{\emptyset,k}^{\infty}$.*

3. *For $k, l \geq 0$, $\phi \in C_{\emptyset,k}^{\infty}$ and $\psi \in C_{\emptyset,l}^{\infty}$ implies that $\phi\psi \in C_{\emptyset,k+l}^{\infty}$.*

4. *If $|\alpha| = k$ then $x^{\alpha} \in C_{\emptyset,k}^{\infty}$.*

5. *If k is a non-negative integer then $|\cdot|^{2k} \in C_{\emptyset,2k}^{\infty}$.*

From Appendix 239 we know that P_n be the space of polynomials of order at most n with complex coefficients, and P be the space of all polynomials with complex coefficients.

Notation 16 ***Fourier transforms of polynomial spaces***

$$\widehat{P}_n = \{\widehat{p} : p \in P_n\}, \qquad \widehat{P} = \{\widehat{p} : p \in P\}.$$

The next theorem proves two relationships between the tempered distributions $\widehat{P_n}$ and the functions $S_{\emptyset,n}$.

Theorem 17 *Suppose that n is a non-negative integer and $u \in S'$. Then using Notation 16:*

1. *$u \in \widehat{P}_n$ iff $[u, \phi] = 0$ for all $\phi \in S_{\emptyset,n}$.*

2. *$u \in \widehat{P}_n$ iff $\phi u = 0$ for all $\phi \in S_{\emptyset,n}$.*

Proof. Part 1 Suppose $[u, \phi] = 0$ for all $\phi \in S_{\emptyset,n}$. This implies that the $\operatorname{supp} u \subset \{0\}$ and by a well known theorem in distribution theory, $u \in \widehat{P}$. Thus $u = \widehat{p}$ for some polynomial p. Suppose $\deg p > n$. For each $|\beta| > n$ choose $\phi_\beta \in S_{\emptyset,n}$ such that $\phi_\beta(x) = x^\beta/\beta!$ in a neighborhood of zero. Then we have $\left(D^\alpha \phi_\beta\right)(0) = \delta_{\alpha,\beta}$ and if the coefficients of p are b_α,
$0 = \left[\widehat{p}, \phi_\beta\right] = \left[p(-D)\left(\phi_\beta\right)\right](0) = (-1)^{|\beta|} b_\beta$, and thus $\deg p < n$ and $u \in \widehat{P}_n$.

Conversely, suppose $u \in \widehat{P}_n$. Then there exists $p \in P_n$ such that $u = \widehat{p} = (2\pi)^{\frac{d}{2}} p(iD)\delta$. Hence, if $\phi \in S_{\emptyset,n}$

$$\begin{aligned} [u, \phi] = \left[(2\pi)^{\frac{d}{2}} p(iD)\delta, \phi\right] = (2\pi)^{\frac{d}{2}} [p(iD)\delta, \phi] = (2\pi)^{\frac{d}{2}} [p(iD)\delta, \phi] &= (2\pi)^{\frac{d}{2}} [\delta, p(-iD)\phi] \\ &= (2\pi)^{\frac{d}{2}} [p(-iD)\phi](0) \\ &= 0. \end{aligned}$$

Part 2 Suppose that $\phi u = 0$ for all $\phi \in S_{\emptyset,n}$. Then for all $\psi \in S$

$$[\phi u, \psi] = [u, \phi\psi] = [\psi u, \phi] = 0,$$

Thus $[\psi u, \phi] = 0$ for all $\phi \in S_{\emptyset,n}$ and so by part 1, $\psi u \in \widehat{P}_n$. Choose a point $x_0 \in \mathbb{R}^d$ and an open ball containing x_0 but not the origin. Choose $\psi \in S$ so that $\psi = 1$ on this ball. Then $u = \psi u \in \widehat{P}_n$ in this neighborhood and so $u = 0$ in this neighborhood. Thus $u = 0$ on $\mathbb{R}^d \backslash 0$. Choose an open ball containing the origin and choose $\psi \in S$ so that $\psi = 1$ on this ball. Then $u = \psi u \in \widehat{P}_n$ in this neighborhood and so $u \in \widehat{P}_n$ in this neighborhood, say $u = \widehat{p}$ where $p \in P_n$. But then $\widehat{p} = 0$ on $\mathbb{R}^d \backslash 0$. Hence $u = \widehat{p}$ on $\mathbb{R}^d$.

Conversely, suppose $u \in \widehat{P}_n$. Then, if $\phi \in S_{\emptyset,n}$ and $\psi \in S$ it follows that $\phi\psi \in S_{\emptyset,n}$. Thus by part 1 of this theorem $[\phi u, \psi] = [u, \phi\psi] = 0$. ■

1.3.1 The functionals $S'_{\emptyset,n}$

The following lemma will enable us to define the space of continuous functionals $S'_{\emptyset,n}$ which in turn will allow us to define the basis distributions in Section 1.6 and the operator $\mathcal{J}$ of Section 1.4.3.

Lemma 18

1. *The space of rapidly decreasing functions S is a locally convex topological space when endowed with the countable seminorm topology defined in Appendix 243.*
2. *A linear functional f defined on a subspace $\mathbb{M}$ of the space S is continuous iff there exists an integer $n \geq 0$ and a constant C such that*
$$|[f, \psi]| \leq C \left\| \sum_{|\alpha| \leq m} (1 + |\cdot|)^m |D^\alpha \psi| \right\|_\infty, \quad \phi \in \mathbb{M}.$$
We write $f \in \mathbb{M}'$.
3. *Any continuous linear functional f on a subspace $\mathbb{M}$ of a locally convex topological vector space $\mathbb{T}$ can be extended non-uniquely to a continuous linear functional f^e on $\mathbb{T}$ ($f^e \in \mathbb{T}'$).*
The set of extensions is $f^e + \mathbb{M}^\perp$ where $\mathbb{M}^\perp$ is the set of annihilators of $\mathbb{M}$ i.e. the members of $\mathbb{T}'$ which are zero on each member of $\mathbb{M}$.

Proof. This lemma can be proved, for example, by using the results and definitions of Chapter V., Volume I of Reed and Simon [17]. ■

Recall Definition 14 of the space $S_{\emptyset,n}$.

Definition 19 ***The spaces*** $S'_{\emptyset,n}$, $n = 1, 2, 3, \ldots$.

$S'_{\emptyset,n}$ is the space of continuous linear functionals on $S_{\emptyset,n}$, where $S_{\emptyset,n}$ has the subspace topology induced by S.

The next theorem gives the criterion we shall use to prove that a linear functional on $S_{\emptyset,n}$ is continuous i.e. is a member of $S'_{\emptyset,n}$. It also supplies the key extension result needed to define the basis distributions.

Theorem 20

1. *A linear functional f on $S_{\emptyset,n}$ is member of $S'_{\emptyset,n}$ iff for some integer $m \geq 0$ and some constant $C > 0$*
$$|[f, \psi]| \leq C \left\| \sum_{|\alpha| \leq m} (1 + |\cdot|)^m |D^\alpha \psi| \right\|_\infty, \quad \phi \in S_{\emptyset,n}.$$
2. *If $f \in S'_{\emptyset,n}$ then there exists $f^e \in S'$ such that $f^e = f$ on $S_{\emptyset,n}$. The set of extensions is $f^e + \widehat{P_n}$. We sometimes say that f can be extended to S as a member f^e of S'.*

Proof. Part 1 follows from part 2 of Lemma 18 with $\mathbb{M} = S_{\emptyset,n}$.

Part 2 We use part 3 of Lemma 18 when $\mathbb{M} = S_{\emptyset,n}$ and $\mathbb{T} = S$. In this case the the set of extensions is $f^e + S_{\emptyset,n}^\perp$ and $g \in S_{\emptyset,n}^\perp$ iff $g \in S'$ and $[g, \phi] = 0$ for all $\phi \in S_{\emptyset,n}$ i.e. $g \in \widehat{P_n}$ by part 1 of Theorem 17. ■

1.4 The function spaces X_w^θ, $\theta = 1, 2, 3, ..$

1.4.1 Introduction

In Section 2 of [11] Light and Wayne introduced the reproducing kernel semi-Hilbert spaces X to formulate the minimal seminorm interpolation problem. The space X was defined using a positive weight function w and had positive order θ. In fact, the spaces X can be described as Beppo-Levi spaces [3] generalized using a positive weight function. In Section 1.4 we extended the class of weight functions to include functions which were positive and continuous except on a set of measure zero. In this section we will use the same definition for our function spaces as Light and Wayne but I will use the notation X_w^θ

for the X spaces of positive order θ generated by the weight function w. Only weight function property W1 is required to define X_w^θ. I will also introduce an alternative definition that I have found quite useful.

By analogy with Sobolev space theory mappings will be constructed between X_w^θ and L^2 in order to prove various properties of X_w^θ. To prove that X_w^θ is a semi-Hilbert space we construct the mappings $\mathcal{I} : X_w^\theta \to L^2$ and $\mathcal{J} : L^2 \to X_w^\theta$ which are adjoints, inverses and isometric isomorphisms in the seminorm sense. Here weight function property W2 is also required. Finally we prove some smoothness results for X_w^θ when the weight function has properties W2 and W3.

Definition 21 ***The semi-inner product spaces*** X_w^θ ***of order*** $\theta = 1, 2, 3, \ldots$.
Suppose w is a weight function i.e. it has property W1. We now define X_w^θ by

$$X_w^\theta = \left\{ f \in S' : \widehat{D^\alpha f} \in L^1_{loc}\left(\mathbb{R}^d\right), \ \int w \left| \widehat{D^\alpha f} \right|^2 < \infty \text{ for all } |\alpha| = \theta \right\}, \tag{1.21}$$

and endow X_w^θ with a semi-inner product and seminorm (part 3 Theorem 25)

$$\langle f, g \rangle_{w,\theta} = \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w \widehat{D^\alpha f}\, \overline{\widehat{D^\alpha g}}, \quad |f|_{w,\theta} = \sqrt{\langle f, f \rangle_{w,\theta}}. \tag{1.22}$$

To prove that X_w^θ is non-empty we need the following lemma:

Lemma 22 *Let $\mathcal{F}$ be a set of points in $\mathbb{R}^d$. Then for any $\eta > 0$ there exists a function f_η such that:*

1. $f_\eta \in C^\infty$,
2. $0 \le f_\eta(x) \le 1$,
3. $f_\eta(x) = 1$ *when* $x \in \mathcal{F}_\eta$,
4. $f_\eta(x) = 0$ *when* $x \notin \mathcal{F}_{3\eta}$.

Here $\mathcal{F}_\eta$ and $\mathcal{F}_{3\eta}$ denote $\eta-$ and $3\eta-$neighborhoods of the set $\mathcal{F}$, as defined in Appendix 241.

Proof. Choose $\eta > 0$. There exists a mollifier $\omega \in C_0^\infty$ satisfying

$$\operatorname{supp} \omega \subseteq B(0;1); \quad \int \omega = 1; \quad \omega \ge 0.$$

Now define the scaled function $\omega_\eta(x) = \eta^{-d} \omega(x/\eta)$ and let $\chi_{2\eta}$ be the characteristic function of the set $\mathcal{F}_{2\eta}$. Then it can be shown that the function $f_\eta(x) = \int \chi_{2\eta}(y)\, \omega_\eta(x-y)\, dy$ has the required properties. ■

Theorem 23 *Suppose the weight function w has property W1 w.r.t. the set $\mathcal{A}$. Then*

$$\left\{ \left(u \left| \cdot \right|^{-\theta} \right)^\vee : u \in C_0^\infty \text{ and } \operatorname{supp} u \subset \mathbb{R}^d \backslash (\mathcal{A} \cup \{0\}) \right\} \subset X_w^\theta, \tag{1.23}$$

and the set on the left is non-empty.

Proof. If Ω is an open set we define $C_0^\infty(\Omega) = \{\phi \in C_0^\infty : \operatorname{supp} \phi \subset \Omega\}$. A weight set is a closed set of measure zero so $\mathcal{B} = \mathcal{A} \cup \{0\}$ is a closed set of measure zero. Then there exists $\eta > 0$ such that the open set $\mathbb{R}^d \backslash \mathcal{B}_\eta$ is non-empty where $\mathcal{B}_\eta$ is the neighborhood set of $\mathcal{B}$ defined by η.

We now use Lemma 22 with $\mathcal{F} = \mathcal{B}$ to show the set on the left of 1.23 is non-empty. Fix $\eta > 0$. Then there exists a function f_η with properties 1 to 4 of Lemma 22. Therefore, $v \in C_0^\infty$ implies $(1 - f_\eta) v \in C_0^\infty$ and $(1 - f_\eta) v = 0$ on $\mathcal{B}_\eta$ i.e. $(1 - f_\eta) v \subset C_0^\infty\left(\mathbb{R}^d \backslash \mathcal{B}\right)$. We conclude that the set on the left of 1.23 is non-empty.

To prove the inclusion 1.23 holds let $f = \left(\frac{u}{|\cdot|^\theta} \right)^\vee$ where $u \in C_0^\infty\left(\mathbb{R}^d \backslash \mathcal{B}\right)$. Because $0 \in \mathcal{B}$ it follows that $\frac{1}{|\cdot|^\theta} \in C^\infty\left(\mathbb{R}^d \backslash \mathcal{B}\right)$ and $\frac{u}{|\cdot|^\theta} \in C_0^\infty\left(\mathbb{R}^d \backslash \mathcal{B}\right)$. Thus $f \in S'$, $\xi^\alpha \widehat{f} = \xi^\alpha \frac{u}{|\cdot|^\theta} \in L^1_{loc}$ and $\int w \left|\cdot\right|^{2\theta} \left| \widehat{f} \right|^2 = \int\limits_{\operatorname{supp} u} w \left|u\right|^2$. But $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ so $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{B}\right)$ and so, $w \left|u\right|^2 = w u \overline{u}$ is continuous on $\operatorname{supp} u$. This means that $\int\limits_{\operatorname{supp} u} w \left|u\right|^2 < \infty$ and so $\int w \left|\cdot\right|^{2\theta} \left| \widehat{f} \right|^2 < \infty$. We conclude that $f \in X_w^\theta$. ■

1.4.2 Some properties of X_w^θ

The space X_w^θ is obviously contained in the space

$$\left\{ f \in S' : \widehat{D^\alpha f} \in L_{loc}^1 \textit{ whenever } |\alpha| = \theta \right\}. \tag{1.24}$$

We will now prove some basic properties of this space of tempered distributions.

Lemma 24 *The functions in the space 1.24 have the properties:*

1. $\widehat{f}(\xi) = \dfrac{(-i)^\theta}{|\xi|^{2\theta}} \sum\limits_{|\alpha|=\theta} \dfrac{\theta!}{\alpha!} \xi^\alpha \widehat{D^\alpha f}(\xi), \quad \xi \neq 0.$
2. $\widehat{f} \in L_{loc}^1 \left(\mathbb{R}^d \backslash 0\right).$
3. $\left|\widehat{f}(\xi)\right| \leq \dfrac{1}{|\xi|^\theta} \sum\limits_{|\alpha|=\theta} \dfrac{\theta!}{\alpha!} \left|\widehat{D^\alpha f}(\xi)\right|, \quad \xi \neq 0.$
4. *Define the function f_F a.e. on $\mathbb{R}^d$ by: $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$.*
   Then $|\cdot|^\theta f_F \in L_{loc}^1$ and $\xi^\alpha f_F \in L_{loc}^1$ when $|\alpha| = \theta$.
5. $\xi^\alpha f_F = \xi^\alpha \widehat{f}$ *a.e.*

Proof. Part 1 is proved by using the identity of part 9 of Appendix 242 to write $|\xi|^{2\theta} \widehat{f}(\xi) = \left(\sum\limits_{|\alpha|=\theta} \dfrac{\theta!}{\alpha!} \xi^{2\alpha} \right) \widehat{f}(\xi)$. **Parts 2 to 4** have easy proofs and these have been omitted. However, some readers may find the definition of the the function f_F in part 4 confusing. The function f_F is **not** a function on $\mathbb{R}^d$ in the distribution sense i.e. it is not an $L_{loc}^1 \left(\mathbb{R}^d\right)$ function. It is a function in the elementary sense which happens to be a member of $L_{loc}^1 \left(\mathbb{R}^d \backslash 0\right)$ and being a member of $L_{loc}^1 \left(\mathbb{R}^d \backslash 0\right)$ defines a function a.e. on $\mathbb{R}^d \backslash 0$ which defines a function a.e. on $\mathbb{R}^d$.

Part 5 True since $\xi^\alpha f_F \in L_{loc}^1$, $\xi^\alpha \widehat{f} \in L_{loc}^1$, $\xi^\alpha f_F = \xi^\alpha \widehat{f}$ on $\mathbb{R}^d \backslash 0$ so $\xi^\alpha f_F = \xi^\alpha \widehat{f}$ a.e. ■

The next theorem derives some properties of X_w^θ and of the function f_F defined in part 4 of the previous lemma. The proof of the formula 1.25 for $|f|_{w,\theta}$ reveals the motivation for using the coefficients $\frac{\theta!}{\alpha!}$ to define the seminorm $|f|_{w,\theta}$. This formula will prove very useful because of its algebraic simplicity. Part 2 of the next theorem provides a way of verifying that a function is in X_w^θ by making use of this formula.

Theorem 25 *Suppose w is a weight function i.e. w has property W1. If $f \in X_w^\theta$ for some integer $\theta \geq 1$ then $\widehat{f} \in L_{loc}^1 \left(\mathbb{R}^d \backslash 0\right)$ and we can a.e. define the function $f_F : \mathbb{R}^d \to \mathbb{C}$ by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Further:*

1. *The seminorm 1.22 satisfies*
$$\int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 = |f|_{w,\theta}^2 . \tag{1.25}$$
2. *An* ***alternative definition*** *of X_w^θ is:*
$$X_w^\theta = \left\{ f \in S' : \xi^\alpha \widehat{f} \in L_{loc}^1 \textit{ if } |\alpha| = \theta; \ \int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 < \infty \right\}. \tag{1.26}$$
   This definition actually makes sense because by part 2 of Corollary 24 the condition $\xi^\alpha \widehat{f} \in L_{loc}^1$ for all $|\alpha| = \theta$ implies that $\widehat{f} \in L_{loc}^1 \left(\mathbb{R}^d \backslash 0\right)$, and so f_F can be defined a.e. on $\mathbb{R}^d$.
3. *The functional $|\cdot|_{w,\theta}$ is a seminorm. In fact* $\mathrm{null}\, |\cdot|_{w,\theta} = P_\theta$ *and we also have $X_w^\theta \cap P = P_\theta$.*

Proof. Part 1 Since f_F is the function defined in part 4 of Lemma 24 we have from part 5 of the same lemma that $\xi^\alpha f_F \in L_{loc}^1$, $\xi^\alpha \widehat{f} \in L_{loc}^1$ and $\xi^\alpha f_F = \xi^\alpha \widehat{f}$ a.e. Hence

$$\begin{aligned}
\int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 = \int w \left(\sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \xi^{2\alpha} \right) \left|f_F\right|^2 = \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w \left|\xi^\alpha f_F\right|^2 &= \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w \left|\xi^\alpha \widehat{f}\right|^2 \\
&= \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w \left|\widehat{D^\alpha f}\right|^2 \\
&= |f|_{w,\theta}^2 < \infty,
\end{aligned}$$

since $f \in X_w^\theta$.

Part 2 Now let $\mathbb{U}$ be the space defined by the right hand side of equation 1.26.

Step 1 Prove that $X_w^\theta \subset \mathbb{U}$: By part 1 of this theorem, if $f \in X_w^\theta$ then $\int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 = \left|f\right|_{w,\theta}^2 < \infty$. Also, $f \in X_w^\theta$ implies $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ and the other requirements for f to be in $\mathbb{U}$ follows from Lemma 24.

Step 2 Prove that $\mathbb{U} \subset X_w^\theta$: Assume $f \in \mathbb{U}$. We need to show that, $\sum\limits_{|a|=\theta} \dfrac{\theta!}{\alpha!} \int w \left|\widehat{D^\alpha f}\right|^2 < \infty$ when $|\alpha| = \theta$. But from the proof of part 1, $\xi^\alpha \widehat{f} = \xi^\alpha f_F \in L_{loc}^1$ and so

$$\sum_{|a|=\theta} \frac{\theta!}{\alpha!} \int w \left|\widehat{D^\alpha f}\right|^2 = \sum_{|a|=\theta} \frac{\theta!}{\alpha!} \int w \left|\xi^\alpha \widehat{f}\right|^2 = \sum_{|a|=\theta} \frac{\theta!}{\alpha!} \int w \left|\xi^\alpha f_F\right|^2 = \sum_{|a|=\theta} \frac{\theta!}{\alpha!} \int w \xi^{2\alpha} \left|f_F\right|^2$$
$$= \int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 < \infty.$$

Part 3 Suppose that $\left|f\right|_{w,\theta} = 0$. Then $f \in S'$, $\widehat{D^\alpha f} \in L_{loc}^1$ for $|\alpha| = \theta$, and so

$$\left|f\right|_{w,\theta}^2 = \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w \left|\widehat{D^\alpha f}\right|^2 = 0.$$

Now $w > 0$ a.e. implies that when $|\alpha| = \theta$, $\widehat{D^\alpha f} = 0$ a.e. and hence $D^\alpha f = 0$ a.e. But by a well known result from distribution theory e.g. corollary to Theorem VI, p.60 Schwartz [20], we can conclude that $f \in P_\theta$. Finally, it is clear that $f \in P_\theta$ implies $\left|f\right|_{w,\theta} = 0$. Thus we have proved null $\left|\cdot\right|_{w,\theta} = P_\theta$.

Clearly $P_\theta \subset X_w^\theta \cap P$. To prove the converse, suppose $f \in X_w^\theta \cap P$ and $f \notin P_\theta$. Then $|\alpha| = \theta$ implies $D^\alpha f \in P \backslash 0$ and $\widehat{D^\alpha f} \notin L_{loc}^1$, but $f \in X_w^\theta$ which implies $\widehat{D^\alpha f} \in L_{loc}^1$, a contradiction. Thus $X_w^\theta \cap P = P_\theta$. ■

The function f_F, introduced in the previous theorem, is of central importance to the theory of this document. We now derive some important properties of this function. To do this we need the following lemma:

Lemma 26 *Suppose the weight function w has property W2. Then for any integer $\theta \geq 1$ there exists a constant $c_{r_2,\theta}$ independent of $\phi \in S_{\emptyset,\theta}$ such that*

$$\left(\int \frac{\left|\phi\right|^2}{w\left|\cdot\right|^{2\theta}}\right)^{1/2} \leq c_{r_2,\theta} \left(\int \frac{1}{w\left|\cdot\right|^{2\lambda(\cdot)}}\right)^{1/2} \sum_{|\alpha|\leq n} \left\|\left(1+\left|\cdot\right|\right)^n D^\alpha \phi\right\|_\infty,$$

where $n = \operatorname{ceil}\{\theta,\sigma\} = \operatorname{ceil}\max\{\theta,\sigma\}$. Here λ is the function introduced in the definition of weight function property W2.

Proof. Suppose r_2 is the parameter in the definition of weight function property W2. Then for $\phi \in S_{\emptyset,\theta}$ we write

$$\begin{aligned}
\int \frac{\left|\phi\right|^2}{w\left|\cdot\right|^{2\theta}} &\leq \int\limits_{|\cdot|\leq r_2} \frac{\left|\phi\right|^2}{w\left|\cdot\right|^{2\theta}} + \int\limits_{|\cdot|\geq r_2} \frac{\left|\phi\right|^2}{w\left|\cdot\right|^{2\theta}} = \int\limits_{|\cdot|\leq r_2} \frac{\left|\phi\right|^2}{\left|\cdot\right|^{2\theta}} \frac{1}{w} + \int\limits_{|\cdot|\geq r_2} \frac{\left|\cdot\right|^{2\sigma}\left|\phi\right|^2}{\left|\cdot\right|^{2\theta}} \frac{1}{w\left|\cdot\right|^{2\sigma}} \\
&\leq \left\|\frac{\phi^2}{\left|\cdot\right|^{2\theta}}\right\|_\infty \int\limits_{|\cdot|\leq r_2} \frac{1}{w} + \max_{|\cdot|\geq r_2}\left(\frac{\left|\cdot\right|^{2\sigma}\left|\phi\right|^2}{\left|\cdot\right|^{2\theta}}\right) \int\limits_{|\cdot|\geq r_2} \frac{1}{w\left|\cdot\right|^{2\sigma}} \\
&\leq \left\|\frac{\phi}{\left|\cdot\right|^{\theta}}\right\|_\infty^2 \int\limits_{|\cdot|\leq r_2} \frac{1}{w} + \left(\max_{|\cdot|\geq r_2}\left(\frac{\left|\cdot\right|^{\sigma}\phi}{\left|\cdot\right|^{\theta}}\right)\right)^2 \int\limits_{|\cdot|\geq r_2} \frac{1}{w\left|\cdot\right|^{2\sigma}} \\
&\leq \left(\max\left\{\left\|\frac{\phi}{\left|\cdot\right|^{\theta}}\right\|_\infty, \max_{|\cdot|\geq r_2}\left(\frac{\left|\cdot\right|^{\sigma}\phi}{\left|\cdot\right|^{\theta}}\right)\right\}\right)^2 \int \frac{1}{w\left|\cdot\right|^{2\lambda(\cdot)}}.
\end{aligned} \tag{1.27}$$

which exists by weight function property W2. Since $n = \text{ceil}\{\theta, \sigma\}$ we can apply the inequality 1.20 to get

$$\left\| \frac{\phi}{|\cdot|^\theta} \right\|_\infty \le \sum_{|\alpha|=\theta} \|D^\alpha \phi\|_\infty \le \sum_{|\alpha|=\theta} \|(1+|\cdot|)^n D^\alpha \phi\|_\infty \le \sum_{|\alpha|\le n} \|(1+|\cdot|)^n D^\alpha \phi\|_\infty .$$

and, since $\phi \in S$

$$\max_{|\cdot|\ge r_2} \left(\frac{|\cdot|^\sigma \phi}{|\cdot|^\theta} \right) \le r_2^{-\theta} \max_{|\cdot|\ge r_2} (|\cdot|^\sigma |\phi|) \le r_2^{-\theta} \|(1+|\cdot|)^n \phi\|_\infty \le r_2^{-\theta} \sum_{|\alpha|\le n} \|(1+|\cdot|)^n D^\alpha \phi\|_\infty .$$

Substituting these inequalities into the right side of 1.27 gives the estimate of this lemma where

$$c_{r_2,\theta} = \max\left\{1, r_2^{-\theta}\right\}. \tag{1.28}$$

Finally, by part 1 of Theorem 20, $\left(\int \frac{|\phi|^2}{w|\cdot|^{2\theta}}\right)^{1/2} \in S'_{\emptyset,\theta}$. ■

Theorem 27 *Suppose $f \in X_w^\theta$. We define the function f_F by: $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$ and $f_F(0) = \{0\}$. Now suppose the weight function w also has property W2. Then f_F has the following properties:*

1. *$f_F = 0$ iff $f \in P_\theta$.*
2. *$f_F \in S'_{\emptyset,\theta}$ with action $\int f_F \phi$, $\phi \in S_{\emptyset,\theta}$. Also $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$.*
3. *For any compact set K*
$$\int_K |\cdot|^\theta f_F \le \left(\int_K \frac{1}{w}\right)^{1/2} |f|_{w,\theta} < \infty.$$
4. *If r_2 is the parameter in the definition of weight function property W2.2*
$$\int_{|\cdot|\ge r_2} \frac{|\cdot|^\theta |f_F|}{(1+|\cdot|)^\sigma} \le \left(\int_{|\cdot|\ge r_2} \frac{1}{w|\cdot|^{2\sigma}}\right)^{1/2} |f|_{w,\theta} < \infty.$$
5. *$|\cdot|^\theta f_F$ is a regular tempered distribution in the sense of Appendix A.5.1. Further, if $|\alpha| = \theta$ then $\xi^\alpha f_F$ is also a regular tempered distribution.*

Proof. Part 1 By parts 1 and 3 of Theorem 25, $f_F = 0$ iff $|f|_{w,\theta} = 0$ iff $f \in P_\theta$.

Part 2 First we show that $\int f_F \phi$ exists for $\phi \in S_{\emptyset,\theta}$. By part 1 of Theorem 25, $\sqrt{w}|\cdot|^\theta f_F \in L^2$ and $|f|_{w,\theta}^2 = \int w|\cdot|^{2\theta} |f_F|^2$, so for $\phi \in S_{\emptyset,\theta}$

$$\left|\int f_F \phi\right| \le \int |f_F \phi| = \int \sqrt{w}|\cdot|^\theta |f_F| \frac{|\phi|}{\sqrt{w}|\cdot|^\theta} \le |f|_{w,\theta} \left(\int \frac{|\phi|^2}{w|\cdot|^{2\theta}}\right)^{1/2},$$

having used the Cauchy-Schwartz inequality. By Lemma 26

$$\left(\int \frac{|\phi|^2}{w|\cdot|^{2\theta}}\right)^{1/2} \le c_{r_2,\theta} \left(\int \frac{1}{w|\cdot|^{2\lambda(\cdot)}}\right)^{1/2} \sum_{|\alpha|\le n} \|(1+|\cdot|)^n D^\alpha \phi\|_\infty ,$$

where $n = \text{ceil}\{\theta, \sigma\}$ and λ is the function defined in weight function property W2. By Theorem 20 this inequality implies $f_F \in S'_{\emptyset,\theta}$.

Now to show that $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$. Since $f_F \in S'_{\emptyset,\theta}$ Theorem 20 implies f_F has a non-unique extension f_F^e to S'. Define $g = (f_F^e)^\vee \in S'$. Next we want to show that $g \in X_w^\theta$. In fact if $|\alpha| = \theta$ and $\phi \in S$ then, since $\xi^\alpha \phi \in S_{\emptyset,\theta}$, $[\xi^\alpha \widehat{g}, \phi] = [\xi^\alpha f_F^e, \phi] = [f_F^e, \xi^\alpha \phi] = [f_F, \xi^\alpha \phi] = [\xi^\alpha f_F, \phi]$ and so $\xi^\alpha \widehat{g} = \xi^\alpha f_F$. But $\xi^\alpha f_F \in L^1_{loc}$ by part 4 of Lemma 24 which means that $\xi^\alpha \widehat{g} \in L^1_{loc}$ and hence by part 2 of Lemma 24 that $\widehat{g} \in L^1_{loc}(\mathbb{R}^d \backslash 0)$. We can now define the function g_F by $g_F = \widehat{g}$ on $\mathbb{R}^d \backslash 0$ and $g_F(0) = 0$. Now if $\phi \in C_0^\infty(\mathbb{R}^d \backslash 0)$ then $\phi \in S_{\emptyset,\theta}$ and it follows that $[g_F, \phi] = [f_F^e, \phi] = [f_F, \phi]$ i.e. $g_F = f_F$ a.e. on $\mathbb{R}^d \backslash 0$ and hence $g_F = f_F$ a.e. on $\mathbb{R}^d$. Thus $|g|_{w,\theta} = |f|_{w,\theta}$ and so $g \in X_w^\theta$. Further $(g - f)_F = 0$ and therefore

$g-f \in P_\theta$ by part 1. Finally, by Theorem 17, $u \in \widehat{P_\theta}$ iff $[u,\phi]=0$ for all $\phi \in S_{\emptyset,\theta}$. Hence $\widehat{g}=\widehat{f}$ on $S_{\emptyset,\theta}$ so that for $\phi \in S_{\emptyset,\theta}$, $\left[\widehat{f},\phi\right]=[\widehat{g},\phi]=[f_F^e,\phi]=[f_F,\phi]$ so that $\widehat{f}=f_F$ on $S_{\emptyset,\theta}$, as required.

Part 3 If K is compact then by using the Cauchy-Schwartz inequality

$$\int\limits_K |\cdot|^\theta |f_F| = \int\limits_K \frac{1}{\sqrt{w}} \sqrt{w} |\cdot|^\theta |f_F| \leq \left(\int\limits_K \frac{1}{w} \right)^{1/2} \left(\int\limits_K w |\cdot|^{2\theta} |f_F|^2 \right)^{1/2} = \left(\int\limits_K \frac{1}{w} \right)^{1/2} |f|_{w,\theta} < \infty,$$

since weight function property W2 implies that $w \in L^1_{loc}$.

Part 4 Using the Cauchy-Schwartz inequality

$$\begin{aligned} \int_{|\cdot|\geq r_2} \frac{|\cdot|^\theta |f_F|}{(1+|\cdot|)^\sigma} \leq \int_{|\cdot|\geq r_2} \frac{|\cdot|^\theta |f_F|}{|\cdot|^\sigma} &= \int_{|\cdot|\geq r_2} \sqrt{w} |\cdot|^\theta |f_F| \frac{1}{\sqrt{w} |\cdot|^\sigma} \\ &\leq \left(\int_{|\cdot|\geq r_2} w |\cdot|^{2\theta} |f_F|^2 \right)^{1/2} \left(\int_{|\cdot|\geq r_2} \frac{1}{w |\cdot|^{2\sigma}} \right)^{1/2} \\ &\leq |f|_{w,\theta} \left(\int_{|\cdot|\geq r_2} \frac{1}{w |\cdot|^{2\sigma}} \right)^{1/2} \\ &< \infty, \end{aligned}$$

since w has property W4.

Part 5 That $|\cdot|^\theta f_F$ is a regular tempered distribution follows immediately from parts 3 and 4. If $|\alpha|=\theta$ then $|\xi^\alpha f_F| \leq |\cdot|^\theta |f_F|$ and so $\xi^\alpha f_F$ is also a regular tempered distribution. ■

Corollary 28 *Suppose the weight function w has property W2, and suppose $f \in S'$, $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and $\int w |\cdot|^{2\theta} |f_F|^2 < \infty$, where f_F was defined in Theorem 27. Then:*

1. *$f_F = 0$ iff $f \in P$.*
2. *$f_F \in S'_{\emptyset,\theta}$ with action $\int f_F \phi$, $\phi \in S_{\emptyset,\theta}$.*
3. *f_F has properties 3 and 4 of Theorem 27.*
4. *f_F has property 5 of Theorem 27 i.e. $|\cdot|^\theta f_F$ is a regular tempered distribution and $\xi^\alpha f_F$ is also a regular tempered distribution when $|\alpha|=\theta$.*

Proof. Part 1 Since $S' \subset \mathcal{D}'$, $f_F = 0$ implies $\widehat{f} \in \mathcal{D}'$ and $\operatorname{supp} \widehat{f} = \{0\}$ which implies $f \in P$. Conversely, $f \in P$ implies $f \in S'$ and $\operatorname{supp} \widehat{f} = \{0\}$ i.e. $f_F = 0$.

Parts 2 to 4 follow from an examination of the proof of Theorem 27. ■

The last corollary allows another definition of the space X^θ_w when the weight function has property W2.

Corollary 29 *Suppose the weight function w has property W2. Then as sets*

$$X^\theta_w = \left\{ f \in S' : \widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right), \int w |\cdot|^{2\theta} |f_F|^2 < \infty, \text{ and } |\alpha|=\theta \text{ implies } \xi^\alpha \widehat{f} = \xi^\alpha f_F \text{ on } S \right\}, \tag{1.29}$$

where $f_F : \mathbb{R}^d \to \mathbb{C}$ is the function defined by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$.

This definition makes sense since by part 4 of Corollary 28 the first two constraints imply that $f_F \in S'_{\emptyset,\theta}$ and when $|\alpha|=\theta$, $\xi^\alpha f_F$ is a regular tempered distribution in the sense of Appendix A.5.1 with action $\int \xi^\alpha f_F \phi$, $\phi \in S$.

Proof. Definition 1.26 of X^θ_w is

$$\left\{ f \in S' : \xi^\alpha \widehat{f} \in L^1_{loc} \ if \ |\alpha|=\theta; \int w |\cdot|^{2\theta} |f_F|^2 < \infty \right\}.$$

Suppose f is a member of the the right side of 1.29. Since w has property W2 part 4 of Corollary 28 implies that when $|\alpha| = \theta$, $\xi^\alpha f_F$ is a regular tempered distribution in the sense of Appendix A.5.1 and so $\xi^\alpha f_F \in L^1_{loc}$. But $\xi^\alpha \widehat{f} = \xi^\alpha f_F$ as distributions so $\xi^\alpha \widehat{f} \in L^1_{loc}$ and thus $f \in X_w^\theta$.

On the other hand, if $f \in X_w^\theta$ then by part 2 of Theorem 25, $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$. Further, by part 2 of Theorem 27, $f_F \in S'_{\emptyset,\theta}$ and $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$. But Theorem 15 implies $\xi^\alpha \psi \in S_{\emptyset,\theta}$ when $|\alpha| = \theta$ so that $\xi^\alpha \widehat{f} = \xi^\alpha f_F$ on S, and we have shown that $f \in X_w^\theta$ implies f is a member of the right side of 1.29. ■

Remark 30 *A word of caution about definition 1.29 of X_w^θ. It is clear that $f \in S'$ and $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ imply that f_F exists. Further, the conditions $f \in S'$, $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and $\int w\, |\cdot|^{2\theta} |f_F|^2 < \infty$ automatically imply that $f_F \in S'_{\emptyset,\theta}$ and $\xi^\alpha f_F \in S'$.*

However, if we prove that $f \in S'$ and $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ so that f_F is defined, and then try to prove that $|\alpha| = \theta$ implies $\xi^\alpha \widehat{f} = \xi^\alpha f_F$ on S, we will first have to show that $\xi^\alpha f_F \in S'$. There is a question of order here.

1.4.3 The operators $\mathcal{I} : X_w^\theta \to L^2$ and $\mathcal{J} : L^2 \to X_w^\theta$

In this section, by analogy with the study of Sobolev spaces, we will define inverse, isometric operators (in the seminorm sense) between X_w^θ and L^2 for $\theta \geq 1$. In the next section these operators will be used to prove the completeness of X_w^θ, again in the seminorm sense, without referring to any other space such as X_w^0. If the weight function also has property W2 these mapping will turn out to be, in the seminorm sense, adjoints, inverses and isometric isomorphisms. It is the properties of these operators that count and the reader can avoid their definitions and the associated lemmas and just study the properties given in Theorems 32, 34 and 35.

Definition 31 ***The operator*** $\mathcal{I} : X_w^\theta \to L^2$, $\theta = 1, 2, 3, \ldots$.

Suppose that the function w has weight function property W1. Using the definition of X_w^θ given in part 2 of Theorem 25 we define the linear operator $\mathcal{I} : X_w^\theta \to L^2$ by

$$\mathcal{I} f = \left(\sqrt{w}\, |\cdot|^\theta f_F\right)^\vee. \tag{1.30}$$

where the function f_F is defined by: $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$ and $f_F(0) = \{0\}$.

Theorem 32 *Properties of $\mathcal{I}$:*

1. *$\mathcal{I}$ is an isometric mapping from X_w^θ to L^2 in the seminorm sense.*

2. $\operatorname{null} \mathcal{I} = P_\theta$.

Proof. Property 1 From part 1 of Theorem 25 we have

$\|\mathcal{I} f\|_2 = \left\| \sqrt{w}\, |\cdot|^\theta f_F \right\|_2 = |f|_{w,\theta}$.

Property 2 Now $\mathcal{I} f = 0$ iff $\sqrt{w}\, |\cdot|^\theta f_F = 0$ iff $f_F = 0$. But $f_F = 0$ iff $f \in P_\theta$ by part 1 of Theorem 27. ■

The next step is to construct an inverse of $\mathcal{I}$, which we will denote by $\mathcal{J}$ and this will require the theory of the spaces $S_{\emptyset,\theta}$ and $S'_{\emptyset,\theta}$ discussed above in Section 1.3. We will be looking for an operator which makes rigorous the formal operator $\left(\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}\right)^\vee$, $g \in L^2$. To justify the inverse-Fourier transform we show that $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta} \in S'_{\emptyset,\theta}$ and then extend $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}$ to S as a member of S'. This operator from L^2 to S' will become our inverse operator $\mathcal{J} : L^2 \to X_w^\theta$. We start by showing $\left(\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}\right)^\vee \in S'_{\emptyset,\theta}$ when $g \in L^2$. Choose $\phi \in S_{\emptyset,\theta}$ and apply the Cauchy-Schwartz inequality to obtain

$$\left| \int \frac{\widehat{g}}{\sqrt{w}\, |\cdot|^\theta} \phi \right| \leq \left(\int |g|^2 \right)^{1/2} \left(\int \frac{|\phi|^2}{w\, |\cdot|^{2\theta}} \right)^{1/2} = \|g\|_2 \left(\int \frac{|\phi|^2}{w\, |\cdot|^{2\theta}} \right)^{1/2}.$$

Now observe that by Lemma 26, $\left(\int \frac{|\phi|^2}{w|\cdot|^{2\theta}}\right)^{1/2} \in S'_{\emptyset,\theta}$ and so $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta} \in S'_{\emptyset,\theta}$. Theorem 20 now allows us to extend the functional $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}$ to S as a member of S'. We can now define the operator $\mathcal{J}$.

Definition 33 ***The operator*** $\mathcal{J} : L^2 \to S'$, $\theta \geq 1$. *Suppose that the weight function w has property W2 and $g \in L^2$. Then $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta} \in S'_{\emptyset,\theta}$. Hence by Theorem 20, this functional can be extended (up to a member of $\widehat{P_\theta}$) to S as an element of S', say f. Now define the class of mappings $\mathcal{J} : L^2 \to S'$ by*

$$\mathcal{J}g = \overset{\vee}{f}.$$

Note that $\mathcal{J}$ is not linear but the next theorem shows it is linear modulo a polynomial in P_θ.

Theorem 34 ***Properties of*** $\mathcal{J}$:

1. $(\mathcal{J}g)_F = \frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}, \quad g \in L^2$.
2. $\mathcal{J} : L^2 \to X_w^\theta$ *and is an isometry in the seminorm sense.*
3. $\mathcal{J}$ *is linear modulo a polynomial in P_θ i.e.*

$$\mathcal{J}\left(\lambda_1 g_1 + \lambda_2 g_2\right) - \lambda_1 \mathcal{J} g_1 - \lambda_2 \mathcal{J} g_2 \in P_\theta.$$

4. $\mathcal{J}g \in P_\theta$ *iff* $g = 0$.

Proof. Properties 1 & 2 To prove $\mathcal{J}g \in X_w^\theta$ we use the definition of X_w^θ given in Corollary 29 i.e. $\mathcal{J}g \in S'$, $\widehat{\mathcal{J}g} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$, and for $(\mathcal{J}g)_F$ defined by: $(\mathcal{J}g)_F = \widehat{\mathcal{J}g}$ on $\mathbb{R}^d \backslash 0$ and $(\mathcal{J}g)_F(0) = \{0\}$, it is required that $\int w|\cdot|^{2\theta} \left|(\mathcal{J}g)_F\right|^2 < \infty$ and $\xi^\alpha \widehat{\mathcal{J}g} = \xi^\alpha (\mathcal{J}g)_F$ on S when $|\alpha| = \theta$.

From the definition of $\mathcal{J}g$ we have $\widehat{\mathcal{J}g} = \frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}$ on $\mathbb{R}^d \backslash 0$ and if $K \subset \mathbb{R}^d \backslash 0$ is compact

$$\int_K \frac{|\widehat{g}|}{\sqrt{w}\,|\cdot|^\theta} \leq \left(\int_K |\widehat{g}|^2\right)^{1/2} \left(\int_K \frac{1}{w\,|\cdot|^{2\theta}}\right)^{1/2} \leq \|g\|_2 \left(\int_K \frac{1}{w\,|\cdot|^{2\theta}}\right)^{1/2}.$$

Since $0 \in 0$ we have $\operatorname{dist}(0; K) > 0$ and so

$$\int_K \frac{1}{w\,|\cdot|^{2\theta}} \leq \frac{1}{(\operatorname{dist}(0; K))^{2\theta}} \int_K \frac{1}{w} < \infty,$$

because $1/w \in L^1_{loc}$ by weight function property W2.1. We now have $(\mathcal{J}g)_F = \frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta}$ which proves property 1. Next

$$|\mathcal{J}g|^2_{w,\theta} = \int w\,|\cdot|^{2\theta} \left|(\mathcal{J}g)_F\right|^2 = \int w\,|\cdot|^{2\theta} \left|\frac{\widehat{g}}{\sqrt{w}\,|\cdot|^\theta}\right|^2 = \|g\|_2^2 < \infty,$$

so that $\mathcal{J}$ is isometric. Now if $|\alpha| = \theta$ and $\psi \in S$ then $\xi^\alpha \psi \in S_{\emptyset,\theta}$ by Theorem 15, and since the operator $\mathcal{J}$ was defined by extending $\frac{\widehat{g}}{\sqrt{w}|\cdot|^\theta} \in S'_{\emptyset,\theta}$ from $S_{\emptyset,\theta}$ to S as a member of S' it follows that

$$\left[\xi^\alpha \widehat{\mathcal{J}g}, \psi\right] = \left[\widehat{\mathcal{J}g}, \xi^\alpha \psi\right] = \left[\frac{\widehat{g}}{\sqrt{w}\,|\cdot|^\theta}, \xi^\alpha \psi\right] = \left[\xi^\alpha (\mathcal{J}g)_F, \psi\right],$$

so that $\xi^\alpha \widehat{\mathcal{J}g} = \xi^\alpha (\mathcal{J}g)_F$ on S, confirming that $\mathcal{J}g \in X_w^\theta$.

Property 3 If $\phi \in S_{\emptyset,\theta}$ then

$$\begin{aligned}\left[\left(\mathcal{J}\left(\lambda_1 g_1 + \lambda_2 g_2\right) - \lambda_1 \mathcal{J} g_1 - \lambda_2 \mathcal{J} g_2\right)^\wedge, \phi\right] &= \left[\frac{\widehat{\lambda_1 g_1 + \lambda_2 g_2}}{\sqrt{w}\,|\cdot|^\theta} - \frac{\widehat{\lambda_1 g_1}}{\sqrt{w}\,|\cdot|^\theta} - \frac{\widehat{\lambda_2 g_2}}{\sqrt{w}\,|\cdot|^\theta}, \phi\right] \\ &= 0.\end{aligned}$$

Part 2 of Theorem 17 gives the required result.

Property 4 The details of this proof are very similar to those of the proof of part 2 of Theorem 32. Suppose $\mathcal{J}g \in P_\theta$. Then $(\mathcal{J}g)_F = 0$, $|\mathcal{J}g|_{w,\theta} = 0$ and thus $\|g\|_2 = 0$ which implies $g = 0$. The argument is easily reversible. ■

Having proved some properties of $\mathcal{I}$ and $\mathcal{J}$ we now study how they interact:

Theorem 35 *Suppose the weight function w has property W2. Then for $\theta \geq 1$ the operators $\mathcal{I} : X_w^\theta \to L^2$ and $\mathcal{J} : L^2 \to X_w^\theta$ interact as follows:*

1. *$(\mathcal{J}\mathcal{I}f)_F = f_F$ when $f \in X_w^\theta$. Also, $\mathcal{J}\mathcal{I} : X_w^\theta \to X_w^\theta$ is an isometry in the seminorm sense.*
2. *$\mathcal{J}\mathcal{I}f - f \in P_\theta$.*
3. *For all choices of $\mathcal{J}$, $\mathcal{I}\mathcal{J} = I$ on L^2.*
4. *$\mathcal{I}$ and $\mathcal{J}$ are adjoints in the sense that $\langle \mathcal{J}g, f \rangle_{w,\theta} = (g, \mathcal{I}f)_2$.*

Proof. Part 1 From part 1 of Theorem 32 $\widehat{\mathcal{I}f} = \sqrt{w} \left|\cdot\right|^\theta f_F$ and from Theorem 34 $(\mathcal{J}g)_F = \frac{\widehat{g}}{\sqrt{w}\left|\cdot\right|^\theta}$. Hence $(\mathcal{J}\mathcal{I}f)_F = \frac{\widehat{\mathcal{I}f}}{\sqrt{w}\left|\cdot\right|^\theta} = f_F$. Further, $\left|\mathcal{J}\mathcal{I}f\right|_{w,\theta}^2 = \int w \left|\cdot\right|^{2\theta} \left|(\mathcal{J}\mathcal{I}f)_F\right|^2 == \int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 = \left|f\right|_{w,\theta}^2$.

Part 2 From part 1, $\mathcal{J}\mathcal{I}f - f \in X_w^\theta$ and $(\mathcal{J}\mathcal{I}f - f)_F = 0$. But by part 1 of Theorem 27 this implies that $\mathcal{J}\mathcal{I}f - f \in P_\theta$.

Part 3 Suppose $g \in L^2$. Then $\mathcal{J}g \in X_w^\theta$ and $(\mathcal{J}g)_F = \frac{\widehat{g}}{\sqrt{w}\left|\cdot\right|^\theta}$. Thus
$\mathcal{I}\mathcal{J}f = \left(\sqrt{w} \left|\cdot\right|^\theta (\mathcal{J}g)_F\right)^\vee = g$.

Part 4

$$\langle \mathcal{J}g, f \rangle_{w,\theta} = \int w \left|\cdot\right|^{2\theta} (\mathcal{J}g)_F \overline{f_F} = \int \sqrt{w} \left|\cdot\right|^\theta \widehat{g} \overline{f_F} = \int \widehat{g} \overline{\sqrt{w} \left|\cdot\right|^\theta f_F} = (g, \mathcal{I}f)_2 .$$

■

1.4.4 The completeness of X_w^θ

In this section, by analogy with Sobolev space theory, we use the operators $\mathcal{I} : X_w^\theta \to L^2$ and $\mathcal{J} : L^2 \to X_w^\theta$ where $\theta \geq 1$, studied in the previous subsection to prove that when w has weight function properties W1 and W2 the semi-inner product space X_w^θ is a semi-Hilbert space i.e. it is complete in the seminorm sense.

Light and Wayne [11] do not define the operator $\mathcal{J} : L^2 \to X_w^\theta$. They only define the operators $\mathcal{I}$ and $\mathcal{J}$ between X_w^0 and L^2 and their proof of the completeness of X_w^θ uses multiple operators and spaces which correspond to each multi-index α such that $|\alpha| = \theta$. In fact, in Definition 2.11 they define the semi-inner product spaces $(Y_\alpha, |f|_\alpha)_{|\alpha|=\theta}$ given by

$$Y_\alpha = \left\{ f \in S' : \widehat{D^\alpha f} \in L_{loc}^1 \text{ and } \int w \left|\widehat{D^\alpha f}\right|^2 < \infty \right\}, \quad |f|_\alpha = \sqrt{\int w \left|\widehat{D^\alpha f}\right|^2},$$

and show that $D^\alpha : Y_\alpha \to X_w^0$ is onto and an isometric isomorphism, which implies each Y_α is complete. Then in Definition 2.14 they define their positive order semi-inner product space $(X, |f|)$ by

$$X = \bigcap_{|\alpha|=\theta} Y_\alpha, \quad |f|^2 = \sum_{|\alpha|=\theta} \frac{1}{\alpha!} \int w \left|\widehat{D^\alpha f}\right|^2,$$

and argue in Theorem 2.15 that the completeness of each Y_α implies that Y is complete.

So there are two ways to prove completeness and although the definition of $\mathcal{J}$ may be difficult, the operators $\mathcal{I}$ and $\mathcal{J}$ have nice properties and it is really only these properties that are important.

Definition 36 ***Completeness in the seminorm sense***

Suppose that $\mathbb{U}$ is a linear space equipped with a seminorm $|\cdot|$. We will refer to $\mathbb{U}$ as being complete in the seminorm sense, if to each Cauchy sequence $\{u_j\} \subset \mathbb{U}$ there corresponds an element $u \in \mathbb{U}$ such that $|u - u_j| \to 0$ as $j \to \infty$. Note that u is no longer uniquely defined by the sequence $\{u_j\}$.

Theorem 37 *Suppose the weight function w only has property W2. Then for $\theta \geq 1$, X_w^θ is complete in the seminorm sense of Definition 36.*

Proof. First note that the operator $\mathcal{J}$ is not linear. Here we use the results of Section 1.4.3 concerning the operators $\mathcal{I}$ and $\mathcal{J}$. Suppose $\{f_k\}$ is Cauchy in X_w^θ. Then $\{\mathcal{I}f_k\}$ is Cauchy in L^2 since $\mathcal{I}$ is an isometry. Since L^2 is complete, $\mathcal{I}f_k \to g$ for some $g \in L^2$. But $\mathcal{J}(\mathcal{I}f_k - g) \in X_w^\theta$ and since $\mathcal{J}$ is an isometry, $|\mathcal{J}(\mathcal{I}f_k - g)|_{w,\theta} = |\mathcal{I}f_k - g|_2 \to 0$.

However, by part 2 of Theorem 35, $\mathcal{J}\mathcal{I}f - f \in P_\theta$. so that

$$|\mathcal{J}(\mathcal{I}f_k - g)|_{w,\theta}^2 = |\mathcal{J}\mathcal{I}f_k - \mathcal{J}g|_{w,\theta}^2 = |f_k - \mathcal{J}g|_{w,\theta}^2,$$

and so $f_k \to \mathcal{J}g$ as $k \to \infty$. ■

The next corollary relates weight function property W3.1 to the completeness of X_w^θ. Here weight function property W3.1 replaces property W2.

Corollary 38 *Suppose the weight function w only has property W3.1 for order $\theta \geq 1$ and $\kappa \geq 0$. Then X_w^θ is complete in the seminorm sense of Definition 36.*

Proof. By part 3 of Theorem 10, property W3.1 implies property W2. Thus the conditions of Theorem 37 are satisfied and X_w^θ is complete. ■

1.4.5 The smoothness of functions in X_w^θ

This theorem provides information about the L^1 behavior of the Fourier transform of functions in X_w^θ, $\theta \geq 1$. This information will be used to prove smoothness properties for X_w^θ when the weight function has properties W3.1 or properties W2.1 and W3.2.

Theorem 39 *If a weight function w has property W2 then given $f \in X_w^\theta$ we can define a.e. the function $f_F : \mathbb{R}^d \to \mathbb{C}$ by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Now we have the following results:*

1. *If w has property W3.1 for order θ and κ, then $\xi^\beta f_F \in L^1$ when $|\beta| \leq \kappa$.*
2. *Suppose w has property W2.1 and also property W3.2 for order θ and κ. Choose $\rho \in C_0^\infty$ such that, $0 \leq \rho \leq 1$, $\rho(x) = 1$ when $|x| \leq r_3$. Then, $(1-\rho)\xi^\beta f_F \in L^1$ when $|\beta| \leq \kappa$.*

Proof. Part 1 First note that by part 2 of Theorem 10 property W3.1 implies property W2 and so f_F is defined. Now by part 1 Theorem 7 property W3.1 implies that for some α satisfying $|\alpha| = \theta$, $\int \frac{|\cdot|^{2\lambda}}{w\,\xi^{2\alpha}} < \infty$ for $0 \leq \lambda \leq \kappa$. So if $|\beta| \leq \kappa$, by using the Cauchy-Schwartz inequality, we obtain

$$\int \left|\xi^\beta f_F\right| \leq \int |\cdot|^{|\beta|}\,|f_F| = \int \frac{|\cdot|^{|\beta|}}{\sqrt{w}\,|\xi^\alpha|}\sqrt{w}\,|\xi^\alpha|\,|f_F| \leq \int \frac{|\cdot|^{|\beta|}}{\sqrt{w}\,|\xi^\alpha|}\sqrt{w}\,|\cdot|^\theta\,|f_F| = \left(\int \frac{|\cdot|^{2|\beta|}}{w\;\xi^{2\alpha}}\right)^{\frac{1}{2}} |f|_{w,\theta},$$

where the last step used part 1 of Theorem 25. The term on the last line is finite since w has property W3.1.

Part 2 First note that, by part 3 of Theorem 10, property W3.2 implies property W2.2. Hence w has property W2 and so f_F is defined.

Since $\widehat{f} \in L^1_{loc}(\mathbb{R}^d\backslash 0)$, $(1-\rho)f_F \in L^1_{loc}$ and $(1-\rho)\xi^\beta f_F \in L^1_{loc}$. Further, since $\rho(\xi) = 1$ when $|\xi| \leq r_3 > 0$, we can use the Cauchy-Schwartz inequality to write

$$\begin{aligned}
\int \left|(1-\rho(\xi))\,\xi^\beta f_F(\xi)\,d\xi\right| &\leq \int\limits_{|\cdot|\geq r_3} |\cdot|^{|\beta|}\,|f_F| = \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^\kappa}{|\cdot|^{(\kappa-|\beta|)}\sqrt{w}\,|\cdot|^\theta}\sqrt{w}\,|\cdot|^\theta\,|f_F| \\
&\leq \frac{1}{(r_3)^{(\kappa-|\beta|)}} \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^\kappa}{\sqrt{w}\,|\cdot|^\theta}\sqrt{w}\,|\cdot|^\theta\,|f_F| \\
&\leq \frac{1}{(r_3)^{(\kappa-|\beta|)}} \left(\int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{2\kappa}}{w\,|\cdot|^{2\theta}}\right)^{\frac{1}{2}} \left(\int\limits_{|\cdot|\geq r_3} w\,|\cdot|^{2\theta}\,|f_F|^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{(r_3)^{(\kappa-|\beta|)}} \left(\int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{2\kappa}}{w\,|\cdot|^{2\theta}}\right)^{\frac{1}{2}} |f|_{w,\theta}.
\end{aligned}$$

Since w has property W3.2 for order θ and κ, the last integral is finite. ■

This theorem corresponds to Light and Wayne's [11] Theorem 2.18 and represents our main smoothness result for X_w^θ spaces.

Theorem 40 *Suppose the weight function w has properties W2.1 and W3 for order θ and κ, as described in Definition 4. Then*

$$X_w^\theta \subset C_{BP}^\infty + C_B^{(\lfloor\kappa\rfloor)} \subset C_{BP}^{(\lfloor\kappa\rfloor)}.$$

Proof. By Theorem 10 properties W2.1 and W3 imply property W2. Now suppose $\rho_0 \in C_0^\infty$, $0 \le \rho_0 \le 1$, and $\rho_0 = 1$ on the ball $B(0;1)$. By Theorem 39, $f \in X_w^\theta$ implies $\xi^\beta (1-\rho_0)\widehat{f} = \xi^\beta (1-\rho_0) f_F \in L^1$ when $|\beta| \le \lfloor\kappa\rfloor$. Hence $D^\beta \left((1-\rho_0)\widehat{f}\right)^\vee \in C_B^{(0)}$ for $|\beta| \le \lfloor\kappa\rfloor$ and so $\left((1-\rho_0)\widehat{f}\right)^\vee \in C_B^{(\lfloor\kappa\rfloor)}$.

We now write $\widehat{f} = \rho_0 \widehat{f} + (1-\rho_0)\widehat{f}$. Taking the inverse Fourier transform of this equation yields

$$f = \left(\rho_0 \widehat{f}\right)^\vee + \left((1-\rho_0)\widehat{f}\right)^\vee.$$

The first term on the right is the inverse Fourier transform of a tempered distribution with compact support and so by Appendix A.6.2, it is a C_{BP}^∞ function. Thus $f \in C_{BP}^\infty + C_B^{(\lfloor\kappa\rfloor)} \subset C_{BP}^{(\lfloor\kappa\rfloor)}$. ■

Corollary 41 *Suppose w is the extended B-spline weight function 1.10 with parameters l and n. Then if w has property W3.1 for θ and some κ it follows that $X_w^\theta \subset C_{BP}^{(n-1+\lfloor\theta/d\rfloor)}$.*

On the other hand, if w has property W3.1 for θ and some κ, it follows that $X_w^\theta \subset C_{BP}^{(n-1+\theta)}$ when $d=1$ and $X_w^\theta \subset C_{BP}^{(n-1)}$ when $d>1$.*

Proof. From 1.17 of Theorem 13, if n and l satisfy property W3.1 for θ and κ then n satisfies $n > \kappa + 1/2 - \lfloor\theta/d\rfloor$. Thus $\kappa = n-1+\lfloor\theta/d\rfloor$ is the largest valid integer value of κ and so Theorem 40 implies $X_w^\theta \subset C_{BP}^{(n-1+\lfloor\theta/d\rfloor)}$.

When w has property W3.1* we employ a similar argument based on Theorem 12. ■

1.5 The function $\frac{1}{w|\cdot|^{2\theta}}$

In this section we will prove some properties of the function $\frac{1}{w|\cdot|^{2\theta}}$, where w is a weight function and θ is a positive integer, which prepare for the definition of the basis distribution and basis function in the next section. The results in this section study the functional $\int \frac{\phi}{w|\cdot|^{2\theta}}$, $\phi \in S_{\emptyset,2\theta}$, and show how it can be extended to a member of the tempered distributions S'. Note that the function $\frac{1}{w|\cdot|^{2\theta}}$ was introduced in Subsection 1.4.3 where it was used to define the operator $\mathcal{J} : L^2 \to X_w^\theta$.

We now prove some properties of the function $\frac{1}{w|\cdot|^{2\theta}}$ where w is a weight function with property W2. The next result shows that $\frac{1}{w|\cdot|^{2\theta}} \in S'_{\emptyset,2\theta}$ and will allow basis distributions to be defined in the next section. Compare this result with a property of $\frac{1}{w|\cdot|^{2\theta}}$ that was proved in Lemma 26.

Theorem 42 *Suppose the weight function w also has property W2. Then the functional $\int \frac{\phi}{w|\cdot|^{2\theta}}$ defined for $\phi \in S_{\emptyset,2\theta}$ is a member of $S'_{\emptyset,2\theta}$. In fact there exists a constant $c_{r_2,\theta}$, independent of $\phi \in S_{\emptyset,2\theta}$, such that*

$$\int \frac{|\phi|}{w\,|\cdot|^{2\theta}} \le (c_{r_2,\theta})^2 \left(\int \frac{1}{w\,|\cdot|^{2\lambda(\cdot)}}\right) \sum_{|\alpha|\le 2n} \left\|(1+|\cdot|)^{2n} D^\alpha \phi\right\|_\infty, \tag{1.31}$$

where $n = \operatorname{ceil}\{\theta,\sigma\}$. Here λ is the function introduced in the definition of weight function property W2, and $c_{r_2,\theta} = \max\left\{1, r_2^{-\theta}\right\}$ is the constant in the estimate of Lemma 26.

Proof. Suppose r_2 is the parameter in the definition of weight function property W2. Then for $\phi \in S_{\emptyset,2\theta}$ we write

$$\begin{aligned}
\int \frac{|\phi|}{w\,|\cdot|^{2\theta}} &= \int\limits_{|\cdot|\le r_2} \frac{|\phi|}{w\,|\cdot|^{2\theta}} + \int\limits_{|\cdot|\ge r_2} \frac{|\phi|}{w\,|\cdot|^{2\theta}} \\
&= \int\limits_{|\cdot|\le r_2} \frac{|\phi|}{|\cdot|^{2\theta}} \frac{1}{w} + \int\limits_{|\cdot|\ge r_2} \frac{|\cdot|^{2\sigma}\,|\phi|}{|\cdot|^{2\theta}} \frac{1}{w\,|\cdot|^{2\sigma}} \\
&\le \left\| \frac{\phi}{|\cdot|^{2\theta}} \right\|_\infty \int\limits_{|\cdot|\le r_2} \frac{1}{w} + \left(\max_{|\cdot|\ge r_2} \left(\frac{|\cdot|^{2\sigma}\,\phi}{|\cdot|^{2\theta}} \right) \right) \int\limits_{|\cdot|\ge r_2} \frac{1}{w\,|\cdot|^{2\sigma}} \\
&\le \max\left\{ \left\| \frac{\phi}{|\cdot|^{2\theta}} \right\|_\infty, \max_{|\cdot|\ge r_2} \left(\frac{|\cdot|^{2\sigma}\,\phi}{|\cdot|^{2\theta}} \right) \right\} \int \frac{1}{w\,|\cdot|^{2\lambda(\cdot)}} \qquad (1.32)
\end{aligned}$$

where the integrals exist by weight function property W2. Since $n = \operatorname{ceil}\{\theta,\sigma\}$ we can apply inequality 1.20 to get

$$\left\| \frac{\phi}{|\cdot|^{2\theta}} \right\|_\infty \le \sum_{|\alpha|=2\theta} \|D^\alpha \phi\|_\infty \le \sum_{|\alpha|=2\theta} \left\| (1+|\cdot|)^{2n} D^\alpha \phi \right\|_\infty \le \sum_{|\alpha|\le 2n} \left\| (1+|\cdot|)^{2n} D^\alpha \phi \right\|_\infty,$$

and, since $\phi \in S$

$$\max_{|\cdot|\ge r_2} \left(\frac{|\cdot|^{2\sigma}\,\phi}{|\cdot|^{2\theta}} \right) \le r_2^{-2\theta} \max_{|\cdot|\ge r_2} \left(|\cdot|^{2\sigma}\,|\phi| \right) \le r_2^{-2\theta} \left\| (1+|\cdot|)^{2n}\,\phi \right\|_\infty \le r_2^{-2\theta} \sum_{|\alpha|\le 2n} \left\| (1+|\cdot|)^{2n} D^\alpha \phi \right\|_\infty .$$

Substituting these inequalities into the right side of 1.32 gives the estimate 1.31 of this lemma. Finally, part 1 of Theorem 20 implies that the functional $\int \frac{\phi}{w|\cdot|^{2\theta}}$ defined on $S_{\emptyset,2\theta}$ is a member of $S'_{\emptyset,2\theta}$. ■

The last theorem now allows us to extend $\int \frac{\phi}{w|\cdot|^{2\theta}}$ to S as a member of S'.

Corollary 43 *Suppose the weight function w has property W2. Then the functional $\int \frac{\phi}{w|\cdot|^{2\theta}}$ of Theorem 42 is a member of $S'_{\emptyset,2\theta}$ and can be extended from $S_{\emptyset,2\theta}$ to S as a member of S', say χ^e. The set of extensions is $\chi^e + \widehat{P}_{2\theta}$.*

Proof. A direct consequence of Theorem 42 and Theorem 20. ■

1.6 Basis distributions and basis functions

It is now time to define the (in general complex-valued) basis distributions that are generated by a weight function with weight property W2. However, it will be shown that if the weight function also has property W3.1 or properties W2.1 and W3.2 then the basis distribution is a continuous function that will be used in the later chapters to construct the basis function interpolants and smoothers studied in those documents. Following Light and Wayne [11] we will define basis distributions *directly* i.e. without reference to the variational problems which define the basis function interpolants and smoothers to be studied in later documents.

We will also calculate the basis functions which correspond to the (tensor product) extended B-spline weight functions introduced in Subsection 1.2.8.

1.6.1 Definition of a basis distribution

We first define a tempered basis distribution of positive order by only assuming weight function properties W1 and W2. Later property W3 will be applied and this ensures the basis distributions are continuous basis functions.

Definition 44 ***Basis distributions and basis functions***

Suppose the weight function w also has property W2. Then by Corollary 43, $\chi = \dfrac{1}{w\left|\cdot\right|^{2\theta}} \in S'_{\emptyset,2\theta}$ for each positive integer θ, and can be extended to S non-uniquely as a member of S' which we denote by χ^e.

*If $G \in S'$ and satisfies $\widehat{G} = \chi^e$, we say G is a **(tempered) basis distribution** of order θ generated by w.*

Thus a basis distribution of order θ is any tempered distribution G which satisfies

$$\left[\widehat{G}, \phi\right] = \int \frac{\phi}{w\left|\cdot\right|^{2\theta}} \; for \; all \; \phi \in S_{\emptyset,2\theta}. \tag{1.33}$$

*For the purposes of this series of documents we will call a basis distribution a **basis function** if it is continuous. We **note** that usually we would call a tempered basis distribution a basis function if it were a regular tempered distribution in the sense of Appendix A.5.1.*

From their definition basis distributions of a given order are not unique. In fact:

Theorem 45 *Suppose the weight function w also has property W2. Suppose G is a basis distribution of order $\theta \geq 1$ generated by w.*

Then the set of basis distributions of order θ is $G + P_{2\theta}$.

Proof. This is a direct consequence of the definition of a basis distribution and Corollary 43. ■

The next theorem will require the following lemma which we give without proof.

Lemma 46 *Suppose $\{x_k\}_{k=1}^{n}$ is a set of distinct points in $\mathbb{R}^d$ and $\{\lambda_k\}_{k=1}^{n}$ is a set of complex numbers. Then the function $f(\xi) = \sum\limits_{j=1}^{n} \lambda_k e^{-ix_k\xi}$ has the properties:*

1. *$f(\xi) = 0$ a.e. implies $\lambda_k = 0$ for all k.*
2. *The null space of f is a closed set of measure zero.*

Theorem 47 *Suppose the points $\{x_k\}_{k=1}^{N}$ are distinct. Then the translated basis distributions $\{G(\cdot - x_k)\}_{k=1}^{N}$ are linearly independent w.r.t. the complex scalars.*

Proof. Suppose $\sum\limits_{k=1}^{n} \lambda_k G(\cdot - x_k) = 0$ and not all $\lambda_k \neq 0$. Then the definition of G implies $\widehat{G} = \frac{1}{w(\xi)|\xi|^{2\theta}}$ on $\mathbb{R}^d \backslash 0$ and so on taking the Fourier transform

$$0 = \sum_{k=1}^{n} \lambda_k e^{-ix_k\xi} \widehat{G} = \left(\sum_{k=1}^{n} \lambda_k e^{-ix_k\xi}\right) \widehat{G} = \frac{\sum_{k=1}^{n} \lambda_k e^{-ix_k\xi}}{w(\xi)\left|\xi\right|^{2\theta}},$$

a.e. on $\mathbb{R}^d \backslash 0$. Hence, since $w(\xi) > 0$ a.e., it follows that $\sum\limits_{k=1}^{n} \lambda_k e^{-i\xi x_k} = 0$ a.e. and on applying Lemma 46 we conclude that $\lambda_k = 0$ for all k. ■

1.6.2 The smoothness of basis distributions; continuous basis functions

In this subsection we look at the smoothness and growth of basis distributions when the weight function has either property W3.1 or properties W2.1 and W3.2 of Definition 4. In the case of property W3.1 we will also derive a simple inverse Fourier transform formula for the basis function and its (bounded) derivatives which only uses L^1 Fourier transform theory. When the weight function has property W3.2 much more effort is required to obtain a 'modified' inverse Fourier transform formula and in Chapter 3 will deal with this.

However, the next theorem will deal with the W3.1 case and it will require a lemma, a standard L^1 inverse Fourier transform result due to Laurent Schwartz. See, for example, Theorem 4.2, p.150 of Malliavin [14].

Lemma 48 *If $f \in S'$ and $\widehat{f} \in L^1$, then $f \in C_B^{(0)}$ and*

$$f(x) = (2\pi)^{-d/2} \int e^{ix\xi} \widehat{f}(\xi)\, d\xi.$$

Theorem 49 *Suppose the weight function w has property* ***W3.1*** *for order θ and smoothness parameter κ. Then $\frac{1}{w|\cdot|^{2\theta}} \in L^1$, $G = \left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee}$ is a basis function of order θ generated by w, and $G \in C_B^{(\lfloor 2\kappa \rfloor)}$. Further, the set of basis distributions is $G + P_{2\theta}$ and*

$$D^{\gamma}G\left(x\right) = \left(2\pi\right)^{-\frac{d}{2}} \int e^{ix\xi} \frac{\left(i\xi\right)^{\gamma}}{w\left(\xi\right)\left|\xi\right|^{2\theta}} d\xi, \quad \left|\gamma\right| \leq \left\lfloor 2\kappa \right\rfloor. \tag{1.34}$$

Proof. By part 2 of Theorem 10 w has property W2 and so the basis distributions are defined. By Theorem 11, $\frac{1}{w|\cdot|^{2\theta}} \in L^1 \subset S'$. Hence $\left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee} \in C_B^{(0)} \subset S'$. Set $G = \left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee}$. Then $G \in S'$ and $\widehat{G} = \frac{1}{w|\cdot|^{2\theta}}$ a.e., implying $G \in C_B^{(0)}$ and so G satisfies

$$\left[\widehat{G}, \phi\right] = \int \frac{\phi}{w\left|\cdot\right|^{2\theta}}, \quad \phi \in S_{\emptyset, 2\theta},$$

and by Definition 44 G is a basis function. By Theorem 45 the set of basis functions is $G + P_{2\theta}$.

Since w has property W3.1, part 1 of Theorem 11 implies that for order θ and κ, $\frac{|\cdot|^{2s}}{w|\cdot|^{2\theta}} \in L^1$ when $0 \leq s \leq \kappa$. Hence

$$\left|\widehat{D^{\gamma}G}\left(\xi\right)\right| = \left|\left(i\xi\right)^{\gamma} \widehat{G}\left(\xi\right)\right| = \frac{\left|\left(i\xi\right)^{\gamma}\right|}{w\left(\xi\right)\left|\xi\right|^{2\theta}} \leq \frac{\left|\xi\right|^{\left|\gamma\right|}}{w\left(\xi\right)\left|\xi\right|^{2\theta}} \in L^1, \quad \left|\gamma\right| \leq \left\lfloor 2\kappa \right\rfloor,$$

and applying Lemma 48 with $f = D^{\gamma}G \in S'$ we obtain 1.34 and $D^{\gamma}G \in C_B^{(0)}$ for $|\gamma| \leq \lfloor 2\kappa \rfloor$ i.e. $G \in C_B^{(\lfloor 2\kappa \rfloor)}$. ∎

Thus when the weight function has property W3.1 for order θ and κ, the basis distribution of order θ must have differentiability of at least $\lfloor 2\kappa \rfloor$ and each derivative is bounded. This is at least twice the minimum smoothness of the functions in X_w^{θ}, which have differentiability of at least $\lfloor \kappa \rfloor$. Continuous basis functions will be used to construct the basis function interpolants and smoothers which are discussed in Chapters 4, 5 and 6.

We will now consider the question of smoothness when a weight function has property W3.2. The next theorem shows that a basis distribution of order θ is a $C_{BP}^{(\lfloor 2\kappa \rfloor)}$ function. This is at least twice the minimum smoothness of the functions in X_w^{θ}, which were shown to have differentiability of at least $\lfloor \kappa \rfloor$.

Theorem 50

1. *If a weight function w has property W2 then $1/w \in S' \cap L^1_{loc}$.*

   Now suppose the weight function w has properties W2.1 and ***W3.2*** *for order θ and κ. Then:*

2. *If G is a basis distribution of order θ then $\left|\cdot\right|^{2\theta} \widehat{G} = \frac{1}{w}$ as tempered distributions.*

3. *Any $f \in S'$ such that $\left|\cdot\right|^{2\theta} \widehat{f} = \frac{1}{w}$ satisfies $f \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$.*

4. *If G is a basis distribution of order θ then $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$.*

Proof. Part 1 Suppose weight function w has property W2. Noting that property W2 is defined by the conditions

$$W2.1 : 1/w \in L^1_{loc}, \quad W2.2 : \int_{|\cdot| \geq r_2} \frac{1}{w\left|\cdot\right|^{2\sigma}} < \infty \ for\ some\ r_2 > 0,$$

we calculate that

$$\left|\int \frac{\phi}{w}\right| \leq 2 \max\left\{\int_{|\cdot| \leq R_2} \frac{1}{w}, \int_{|\cdot| \leq R_2} \frac{1}{w\left|\cdot\right|^{2\sigma}}\right\} \left\|\left(1 + \left|\cdot\right|\right)^{2\lceil \sigma \rceil} \phi\right\|_{\infty}, \quad \phi \in S,$$

where $\lceil \sigma \rceil$ denotes the *ceiling* of σ. Hence $1/w \in S'$, since $\left\|\left(1 + \left|\cdot\right|\right)^{2\lceil \sigma \rceil} \phi\right\|_{\infty}$ is one of the seminorms used in Definition 243 to specify the topology of S'.

Part 2 By part 4 of Theorem 10, w has property W2 and so the basis distribution is defined and by part 1, $1/w \in S'$. If $\phi \in S$ then by Theorem 15, $|\cdot|^{2\theta} \phi \in S_{\emptyset,2\theta}$ and so by 1.33 of the basis distribution definition

$$\left[|\cdot|^{2\theta} \widehat{G}, \phi\right] = \left[\widehat{G}, |\cdot|^{2\theta} \phi\right] = \left[\frac{1}{|\cdot|^{2\theta} w}, |\cdot|^{2\theta} \phi\right] = \int \frac{1}{|\cdot|^{2\theta} w} |\cdot|^{2\theta} \phi = \int \frac{1}{w}\phi = \left[\frac{1}{w}, \phi\right].$$

Part 3 In part 2 it was shown that $1/w \in S'$ and that there exists $f \in S'$ such that $|\cdot|^{2\theta} \widehat{f} = \frac{1}{w}$. The rest of this part is based on the fact that $g \in L^1$ implies $\widehat{g} \in C_B^{(0)}$. Since $r_3 > 0$ there exists a function $\psi \in C_0^\infty$ such that $0 \le \psi \le 1$, $\psi = 1$ on $B(0; r_3)$ and $\psi = 0$ outside $B(0; 2r_3)$. Hence $\dfrac{1-\psi}{|\cdot|^{2\theta}} \in L^1_{loc} \cap C_B^\infty$ and since $|\cdot|^{2\theta} \widehat{f} = \frac{1}{w} \in S'$ we have

$$\frac{1-\psi}{|\cdot|^{2\theta}} |\cdot|^{2\theta} \widehat{f} = (1-\psi)\widehat{f} = \frac{1-\psi}{w|\cdot|^{2\theta}} \in S' \cap L^1_{loc}. \tag{1.35}$$

The next step is to prove that $x^\alpha (1-\psi) \widehat{f} \in L^1$ when $|\alpha| \le \lfloor 2\kappa \rfloor$. But

$$x^\alpha (1-\psi)\widehat{f} = x^\alpha \frac{1-\psi}{w|\cdot|^{2\theta}}, \tag{1.36}$$

and if $|\alpha| \le \lfloor 2\kappa \rfloor$

$$\int \left| x^\alpha \frac{1-\psi}{w|\cdot|^{2\theta}} \right| dx \le \int |\cdot|^{|\alpha|} \frac{1-\psi}{w|\cdot|^{2\theta}} = \int\limits_{|\cdot| \ge r_3} |\cdot|^{|\alpha|} \frac{1-\psi}{w|\cdot|^{2\theta}} \le \int\limits_{|\cdot| \ge r_3} \frac{|\cdot|^{|\alpha|}}{w|\cdot|^{2\theta}},$$

and

$$\int\limits_{|\cdot| \ge r_3} \frac{|\cdot|^{|\alpha|}}{w|\cdot|^{2\theta}} = \int\limits_{|\cdot| \ge r_3} \frac{|\cdot|^{2\kappa}}{|\cdot|^{2\kappa-|\alpha|} w|\cdot|^{2\theta}} \le \frac{1}{(r_3)^{2\kappa-|\alpha|}} \int\limits_{|\cdot| \ge r_3} \frac{|\cdot|^{2\kappa}}{w|\cdot|^{2\theta}} < \infty,$$

by property W3.2. Hence $x^\alpha \dfrac{1-\psi}{w|\cdot|^{2\theta}} \in L^1$ when $|\alpha| \le \lfloor 2\kappa \rfloor$, and so $D^\alpha \left(\dfrac{1-\psi}{w|\cdot|^{2\theta}}\right)^\vee \in C_B^{(0)}$ when $|\alpha| \le \lfloor 2\kappa \rfloor$ i.e.

$$\left(\frac{1-\psi}{w|\cdot|^{2\theta}}\right)^\vee \in C_B^{(\lfloor 2\kappa \rfloor)}. \tag{1.37}$$

Also, from equation 1.35

$$\left(\frac{1-\psi}{w|\cdot|^{2\theta}}\right)^\vee = \left((1-\psi)\widehat{f}\right)^\vee = f - \left(\psi\widehat{f}\right)^\vee,$$

or on rearranging

$$f = \left(\frac{1-\psi}{w|\cdot|^{2\theta}}\right)^\vee + \left(\psi\widehat{f}\right)^\vee. \tag{1.38}$$

We already know from 1.37 that the first term on the right is a $C_B^{(\lfloor 2\kappa \rfloor)}$ function and since ψ has bounded support $\psi\widehat{f}$ is a distribution with compact support and by Appendix A.6.2 the inverse Fourier transform is a C_{BP}^∞ function. Thus $f \in C_B^{(\lfloor 2\kappa \rfloor)} + C_{BP}^\infty \subset C_{BP}^{(\lfloor 2\kappa \rfloor)}$.

Part 4 Follows directly from parts 2 and 3. ■

We summarize our basis function smoothness results in:

Summary 51 *Suppose a weight function w has properties W2 and W3 for order θ and smoothness parameter κ. Suppose G is a basis distribution of order θ generated by w. Then:*

1. *If w has property W3.1 then $G \in C_B^{(\lfloor 2\kappa \rfloor)}$ (Theorem 49).*

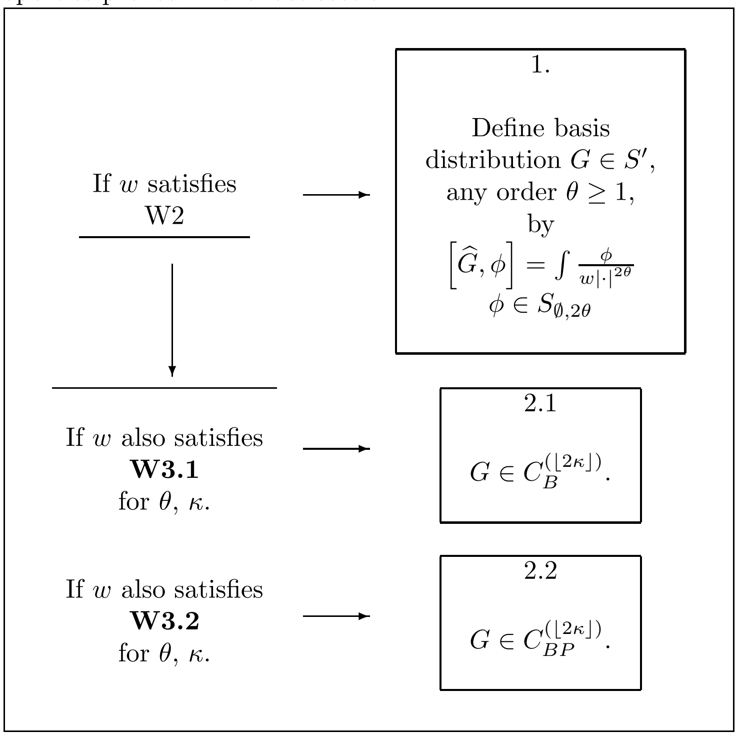
2. If w has property W3.2 then $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ (part 4 Theorem 50).

Remark 52 *Note that if a weight function has property W2 then basis distributions are defined for all positive orders. However, if the weight function also has property W3 we have only proved that the basis distribution is continuous* ***if the order of the basis distribution matches the order of the weight function****.*

Continuous basis distributions can be used to express the solution of the interpolation and smoothing problems discussed later.

1.6.3 Summary diagram

Figure 1.6.3 illustrates the relationships between the weight function properties and the basis distribution properties proved in the last section.



Box 1 A tempered basis distribution G of order $\theta \geq 1$ is defined for weight functions w which satisfy property W2. When θ is positive this coincides with Light and Wayne's definition.

Box 2.1 Suppose w also has the property W3.1 for order $\theta \geq 1$ and smoothness parameter κ. Then by Theorem 49 the basis distribution G of order θ is a basis function in $C_B^{(\lfloor 2\kappa \rfloor)}$.

Box 2.2 Suppose w also has the property W3.2 for order $\theta \geq 1$ and smoothness parameter $\kappa \geq 0$. Then by Theorem 50 the basis distribution G of order θ is a basis function in $C_{BP}^{(\lfloor 2\kappa \rfloor)}$. There are no growth estimates near infinity.

1.6.4 The extended B-spline basis functions

Suppose w is an extended B-spline weight function 1.10 which has property W3.1 for order θ and smoothness parameter κ. Then by Theorem 49, $\left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee}$ is a basis function of order θ. However the formula $\left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee}$ does not in general lend itself to the calculation of a 'closed form' suitable for numeric calculations involving basis functions e.g. basis function smoothers, so we will derive an alternative convolution form for the basis functions in Theorem 56**. In fact these basis functions turn out to be derivatives of the zero order extended B-spline basis function convolved with a thin-plate spline basis function**. To show this we will need three lemmas. The first lemma gives some properties of the zero order extended B-splines proved in Chapter 1 of Williams [22].

Lemma 53 *Suppose w is the extended B-spline weight function 1.10 and suppose $1 \le n \le l$. Then $1/w \in L^1$ and the basis function $G_0 = \left(\frac{1}{w}\right)^{\vee}$ of* ***order zero*** *generated by w is the tensor product $G_0(x) = \prod\limits_{k=1}^{d} H(x_k)$ where*

$$H(t) = (-1)^{l-n} \frac{(2\pi)^{l/2}}{2^{2(l-n)+1}} D^{2(l-n)} \left((*\Lambda)^l \right) \left(\tfrac{t}{2}\right), \quad t \in \mathbb{R}^1,$$

and $(*\Lambda)^l$ denotes the convolution of l 1-dimensional hat functions.

Further, if $n < l$ we have

$$D^{2(l-n)} (*\Lambda)^l = (-1)^{l-n} (*\Lambda)^{l-n} * \sum_{k=-(l-n)}^{l-n} (-1)^{|k|} \binom{l-n}{|k|} \Lambda(\cdot - k),$$

$H \in C_0^{(2n-2)}\left(\mathbb{R}^1\right)$, $D^{2n-1}H$ is a bounded, piecewise constant function and $D^{2n}H$ is the sum of a finite number of translated delta functions.

Finally $G_0 \in C_0^{(2n-2)}\left(\mathbb{R}^d\right)$, the derivatives $\{D^\alpha G_0\}_{|\alpha|=2n-1}$ are bounded functions with bounded support, and the derivatives $\{D^\alpha G_0\}_{|\alpha|=2n}$ are the tensor product of bounded functions with bounded support and finite sums of translated delta functions.

Proof. The zero order weight and basis functions are defined above in Subsection 1.2.1. By Theorem 7 of Williams [22] w has the zero order property W2 (equation 1.1) for $\kappa = n-1$. Thus $1/w \in L^1$ by property W2 and from 1.2 the zero order basis function is $G_0 = \left(\frac{1}{w}\right)^{\vee}$. This lemma now follows from Theorem 29 of Williams [22]. ∎

Lemma 54 *If w is the extended B-spline weight function 1.10 and $1 \le n < l$ then $\frac{1}{w} \in C^\infty_{\emptyset,2d(l-n)} \cap C^\infty_{BP}$.*

Proof. G_0 is a distribution and has bounded support so by Appendix A.6.2, $\widehat{G_0} = \frac{1}{w} \in C^\infty_{BP}$. Further, if $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$ we write

$$\begin{aligned} \frac{1}{w(x)} = (2\pi)^{\frac{d}{2}} \prod_{i=1}^{d} \frac{\sin^{2l} x_i}{x_i^{2n}} = (2\pi)^{\frac{d}{2}} \prod_{i=1}^{d} x_i^{2(l-n)} \frac{\sin^{2l} x_i}{x_i^{2l}} &= (2\pi)^{\frac{d}{2}} x^{2(l-n)\mathbf{1}} \prod_{i=1}^{d} \frac{\sin^{2l} x_i}{x_i^{2l}} \\ &= (2\pi)^{\frac{d}{2}} x^{2(l-n)\mathbf{1}} \left(\prod_{i=1}^{d} \frac{\sin^2 x_i}{x_i^2} \right)^l \\ &= (2\pi)^{\frac{d}{2}} x^{2(l-n)\mathbf{1}} \left((2\pi)^{\frac{d}{2}} \widehat{\Lambda}(2x) \right)^l, \end{aligned}$$

where the last step used property 1.5 of the tensor product hat function defined by 1.3 and 1.4. But by 1.3 Λ has bounded support so by Appendix A.6.2, $\widehat{\Lambda} \in C^\infty_{BP}$. Theorem 15 can now be applied to prove that $x^{2(l-n)\mathbf{1}} \in C^\infty_{\emptyset,2d(l-n)} \cap C^\infty_{BP}$ and consequently that $\frac{1}{w} \in C^\infty_{\emptyset,2d(l-n)} \cap C^\infty_{BP}$. ∎

Lemma 55 *Suppose T_m is the thin-plate spline basis function with integer order $m > d/2$ i.e.*

$$T_m(x) = \begin{cases} (-1)^{m-(d-2)/2} |x|^{2m-d} \log|x|, & d \text{ even}, \\ (-1)^{m-(d-1)/2} |x|^{2m-d}, & d \text{ odd}. \end{cases} \tag{1.39}$$

Then:

1. *For some constant $c_m > 0$*
$$\widehat{T_m} = \frac{c_m}{|\cdot|^{2m}} \text{ on } S_{\emptyset,2m}. \tag{1.40}$$
2. *$T_m \in C_{BP}^{(2m-d-1)} \cap C^\infty\left(\mathbb{R}^d \backslash 0\right)$.*
3. *When d is odd, $D^\alpha T_m$ is a bounded function when $|\alpha| = 2m-d$.*
4. *When d is even, $|D^\alpha T_m(x)| \le k_\alpha (1 + |\log|x||)$ when $|\alpha| = 2m-d$.*

Proof. Part 1 The thin-plate (surface) spline basis functions T_m given by 1.39 are those of Dyn [4] equation 4 multiplied by the power of minus one which makes its Fourier transform 1.40 a positive function outside the origin. The power of minus one can be determined by comparing the Fourier transform of the differential equation 7 of Dyn with equations 23 and 24. Alternatively, Subsection 2.3.1 of Chapter 2 gives the relevant formulas based on Dyn's paper e.g. $c_m = e(m - d/2)$ where the function e is defined in equation 2.5.

The modified form of Dyn's equation 7 is

$$|D|^{2m} T_m = (-1)^m (2\pi)^{d/2} c_m \delta,$$

for some constant $c_m > 0$. Taking the Fourier transform of this equation yields $(-1)^m |\cdot|^{2m} \widehat{T_m} = (-1)^m (2\pi)^{d/2} c_m \widehat{\delta} = (-1)^m c_m$ i.e.

$$|\cdot|^{2m} \widehat{T_m} = c_m, \tag{1.41}$$

and thus

$$\widehat{T_m} = \frac{c_m}{|\cdot|^{2m}} \text{ on } \mathbb{R}^d \backslash 0. \tag{1.42}$$

To prove 1.40 we note that the constant function $\frac{1}{c_m}$ has weight function properties W1, W2 and W3.2 for order m and κ iff $m > \kappa + d/2$. Hence by Definition 44 every basis function G of order m generated by the weight function $\frac{1}{c_m}$ satisfies $\widehat{G} = \frac{c_m}{|\cdot|^{2m}}$ on $S_{\emptyset,2m}$ and so $\widehat{G} = \frac{c_m}{|\cdot|^{2m}}$ on $\mathbb{R}^d \backslash 0$ as distributions. Equation 1.42 now implies $\operatorname{supp}\left(\widehat{T_m} - \widehat{G}\right) \subset \{0\}$ so that $\widehat{T_m} - \widehat{G}$ is a finite sum of derivatives of delta functions and hence $T_m - G = p$ for some polynomial p.

But by Theorem 15, $|\cdot|^{2m} \in C^\infty_{\emptyset,2m} \cap C^\infty_{BP}$ and $\phi \in S$ implies $|\cdot|^{2m} \phi \in S_{\emptyset,2m}$. Hence $\left[|\cdot|^{2m} \widehat{G}, \phi\right] = \left[\widehat{G}, |\cdot|^{2m} \phi\right] = \left[\frac{c_m}{|\cdot|^{2m}}, |\cdot|^{2m} \phi\right] = [c_m, \phi]$ and $|\cdot|^{2m} \widehat{G} = c_m$. Consequently $0 = |\xi|^{2m} \left(\widehat{T_m} - \widehat{G}\right) = |\xi|^{2m} \widehat{p}$ which implies that $p \in P_{2m}$ and therefore $T_m - G \in P_{2m}$. Theorem 45 now shows that T_m is a basis function of order m with weight function $\frac{1}{c_m}$ and 1.40 follows.

Part 2 Clearly $T_m \in C^\infty\left(\mathbb{R}^d \backslash 0\right)$. Now $m > \kappa + d/2$ so that $2m > 2\kappa + d$ and $\lfloor 2\kappa \rfloor = 2m - d - 1$ and thus $T_m \in C^{(2m-d-1)}_{BP}$ by Theorem 48.

Part 3 When d is odd, $T_m(x) = (-1)^{m-(d-1)/2} |x|^{2m-d}$. Thus T_m is a homogeneous function of order $2m - d$ and so $D^\alpha T_m(\lambda x) = \lambda^{2m-d} D^\alpha T_m(x) = \lambda^{|\alpha|} (D^\alpha T_m)(x)$ which implies $D^\alpha T_m(x) = |x|^{2m-d-|\alpha|} (D^\alpha T_m)\left(\frac{x}{|x|}\right)$. When $|\alpha| = 2m - d$, $D^\alpha T_m(x) = (D^\alpha T_m)\left(\frac{x}{|x|}\right)$ and $D^\alpha T_m$ is bounded.

Part 4 When d is even, $T_m(x) = (-1)^{m-(d-2)/2} |x|^{2m-d} \log|x|$. Hint: First observe that $D_k \log|x|$ is a homogeneous function of order -1 and that $|x|^{2n}$ is a homogeneous function of order $2n$. Then, on setting $n = (2m - d)/2$, expand $D^\alpha \left(|x|^{2n} \log|x|\right)$ using the Leibniz rule A.1. ■

Theorem 56 *Suppose w is an extended B-spline weight function 1.10 which has property W3.1 for order θ and κ. Then $G_0 = \left(\frac{1}{w}\right)^\vee$ is the zero order extended B-spline basis function and $G_\theta = \left(\frac{1}{w|\cdot|^{2\theta}}\right)^\vee$ is an order θ basis function. In fact, the basis function G_θ has the form*

$$G_\theta = \begin{cases} \frac{1}{c_\theta} G_0 * T_\theta, & 2\theta > d, \\ \frac{(-1)^{\frac{d+1}{2}-\theta}}{c_{\frac{d+1}{2}}} |D|^{d+1-2\theta} G_0 * T_{\frac{d+1}{2}}, & 2\theta < d,\ d \text{ odd}, \\ \frac{(-1)^{\frac{d+2}{2}-\theta}}{c_{\frac{d+2}{2}}} |D|^{d+2-2\theta} G_0 * T_{\frac{d+2}{2}}, & 2\theta \le d,\ d \text{ even}, \end{cases} \tag{1.43}$$

where T_m is the thin-plate spline basis function of integer order $m > d/2$ given in Lemma 55. The convolutions of 1.43 can be written

$$(G_0 * T_m)(x) = (2\pi)^{-d/2} \int G_0(y) T_m(x - y)\, dy, \quad m > d/2, \tag{1.44}$$

with $G_0 * T_m$ being a regular tempered distribution as defined in Appendix A.5.1.
Also, $\max \lfloor 2\kappa \rfloor = 2n - 2 + 2\lfloor \theta/d \rfloor$ and $G_\theta \in C_B^{(2n-2+2\lfloor \theta/d \rfloor)}$.

Proof. Since w has property W3.1 for parameters θ and κ it follows from Lemma 53 that $\frac{1}{w} \in L^1$ and that $G = \left(\frac{1}{w}\right)^{\vee}$ is the unique zero order basis function. Also, from Theorem 49 we have that $\left(\frac{1}{w|\cdot|^{2\theta}}\right)^{\vee}$ is a basis function of order θ. Now suppose that $\phi \in S$. By Lemma 54 $\frac{1}{w} \in C^{\infty}_{\emptyset,2d\theta} \cap C^{\infty}_{BP}$, and so

$$\frac{1}{w}\phi \in S_{\emptyset,2\theta}. \tag{1.45}$$

$\boxed{\textbf{Case } 2\theta > d}$ By Lemma 55, $\frac{1}{c_{2\theta}}\widehat{T_\theta} = \frac{1}{|\cdot|^{2\theta}}$ on $S_{\emptyset,2\theta}$, and thus

$$\begin{aligned}\left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{w\,|\cdot|^{2\theta}}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta}}, \frac{1}{w}\phi\right] = \left[\frac{1}{c_{2\theta}}\widehat{T_\theta}, \frac{1}{w}\phi\right] &= \left[\frac{1}{c_{2\theta}}\frac{1}{w}\widehat{T_\theta}, \phi\right] \\ &= \left[\frac{1}{c_{2\theta}}\widehat{G_0}\widehat{T_\theta}, \phi\right] \\ &= \left[\frac{1}{c_{2\theta}}\widehat{G_0 * T_\theta}, \phi\right],\end{aligned}$$

so that $G_\theta = \frac{1}{c_{2\theta}} G_0 * T_\theta$.

$\boxed{\textbf{Case } 2\theta < d \text{ and } d \text{ is odd}}$ By Lemma 55, $\widehat{T_{\frac{d+1}{2}}} = c_{\frac{d+1}{2}} \frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}$ on $S_{\emptyset,d+1}$, and thus 1.45 implies

$$\begin{aligned}\left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{w\,|\cdot|^{2\theta}}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta}}, \frac{\phi}{w}\right] &= \left[\frac{|\cdot|^{d+1-2\theta}}{|\cdot|^{d+1}}, \frac{\phi}{w}\right] \\ &= \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \frac{|\cdot|^{d+1-2\theta}}{w}\phi\right] \\ &= \left[\frac{1}{c_{\frac{d+1}{2}}}\widehat{T_{\frac{d+1}{2}}}, \frac{|\cdot|^{d+1-2\theta}}{w}\phi\right] \\ &= \left[\frac{1}{c_{\frac{d+1}{2}}}\frac{|\cdot|^{d+1-2\theta}}{w}\widehat{T_{\frac{d+1}{2}}}, \phi\right] \\ &= \left[\frac{1}{c_{\frac{d+1}{2}}}|\cdot|^{d+1-2\theta}\,\widehat{G_0}\;\widehat{T_{\frac{d+1}{2}}}, \phi\right] \\ &= \left[\frac{1}{c_{\frac{d+1}{2}}}|\cdot|^{d+1-2\theta}\left(G_0 * T_{\frac{d+1}{2}}\right)^{\wedge}, \phi\right] \\ &= \left[\frac{(-1)^{\frac{d+1}{2}-\theta}}{c_{\frac{d+1}{2}}}\left(|D|^{d+1-2\theta}\,G_0 * T_{\frac{d+1}{2}}\right)^{\wedge}, \phi\right],\end{aligned}$$

so that $G_\theta = \frac{(-1)^{\frac{d+1}{2}-\theta}}{c_{\frac{d+1}{2}}}|D|^{d+1-2\theta}\,G_0 * T_{\frac{d+1}{2}}$.

$\boxed{\textbf{Case } 2\theta \le d \text{ and } d \text{ even}}$ In a very similar manner to the previous case we can obtain $G_\theta = \frac{(-1)^{\frac{d+2}{2}-\theta}}{c_{\frac{d+2}{2}}}|D|^{d+2-2\theta}\,G_0 * T_{\frac{d+2}{2}}$. The last three cases combine to prove 1.43.

Since Theorem 53 implies $G_0 \in C_0^{(2n-2)}\left(\mathbb{R}^d\right)$ and Lemma 55 shows that $T_m \in C_{BP}^{(2m-d-1)}$, 1.44 follows from the convolution formulas of part 5 of Appendix 245.

Finally, if a weight function w has property W3.1 for order θ and κ then Theorem 49 implies $G_\theta \in C_B^{(\lfloor 2\kappa \rfloor)}$. But from 1.17 the largest valid integer value of 2κ is $2n - 2 + 2\lfloor \theta/d \rfloor$ so it follows that $G_\theta \in C_B^{(2n-2+2\lfloor \theta/d \rfloor)}$. ■

1.6.5 *Convolution formulas for basis functions generated by weight functions with property W3.1*

In this subsection we will develop convolution formulas designed to facilitate the expression of positive order basis functions in a closed form which can be used for numerical work.

Lemma 57 *Suppose the weight function w has property W3.1 for θ and κ. Then there exists a multi-index α such that $|\alpha| = 2\theta$ and:*

1. $w_\alpha(\xi) = \xi^\alpha w(\xi)$ *is a zero order weight function with property W2 for parameter κ.*
2. $G_\alpha = (1/w_\alpha)^\vee \in C_B^{(\lfloor 2\kappa \rfloor)}$ *is the unique zero order basis function generated by w_α.*
3. $G_0 = (1/w)^\vee \in S'$ *and we call this the zero order distribution generated by w.*
4. $G_\theta = \left(\frac{1}{w|\cdot|^{2\theta}}\right)^\vee \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ *is a basis function of order θ generated by w.*

Proof. The properties W1 and W2 of a zero order weight function are given in Subsection 1.2.1. Observe that the properties W1 for both the zero and positive order cases are identical.

Part 1 Since w has property W1 it has property W1 as a zero order weight function and must be continuous and positive outside some closed set $\mathcal{A}$ with measure zero. Hence w_α is continuous and positive outside the closed set $\mathcal{A} \cup \bigcup_{k=1}^{d} \{x : x_k = 0\}$. Since this set also has measure zero, w_α also has property W1 i.e. it is a zero order weight function.

By Theorem 7 $\int \frac{|x|^{2t}}{w(x)\, x^{2\alpha}} dx < \infty$ when $0 \le t \le \kappa$. Thus $\int \frac{|x|^{2t}}{w_\alpha(x)} dx < \infty$ when $0 \le t \le \kappa$ and so has property W2 for parameter κ as a zero order weight function.

Part 2 From Subsection 1.2.1 the zero order basis function is $(1/w_\alpha)^\vee$ and it is a function in $C_B^{(\lfloor 2\kappa \rfloor)}$.

Part 3 By part 1 of Theorem 50, $1/w \in S'$.

Part 4 By Theorem 49, $G_\theta = \left(\frac{1}{w|\cdot|^{2\theta}}\right)^\vee \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ is a basis function of order θ generated by w. ∎

We begin with the case where G_α has bounded support, which simplifies the convolution.

Theorem 58 *Suppose the weight function w has property W3.1 for order θ and κ. Assume the weight function w_α, the basis functions G_α, G_θ and the basis distribution G_0 are as defined above in Lemma 57. Then if G_α has bounded support the convolution formulas*

$$G_\theta = \begin{cases} \frac{(-1)^\theta}{c_\theta} D^{2\alpha} G_\alpha * T_\theta, & 2\theta > d, \\ \frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} |D|^{d+1-2\theta} D^{2\alpha} G_\alpha * T_{\frac{d+1}{2}}, & 2\theta < d,\ d \text{ odd}, \\ \frac{(-1)^{\frac{d+2}{2}}}{c_{\frac{d+2}{2}}} |D|^{d+2-2\theta} D^{2\alpha} G_\alpha * T_{\frac{d+2}{2}}, & 2\theta \le d,\ d \text{ even}, \end{cases} \tag{1.46}$$

hold. Here T_m is the thin-plate spline basis function of integer order $m > d/2$ defined in Lemma 55. The convolutions of 1.46 can be written

$$(G_\alpha * T_m)(x) = (2\pi)^{-d/2} \int G_\alpha(y)\, T_m(x-y)\, dy, \quad m > d/2, \tag{1.47}$$

with $G_\alpha * T_m$ being a regular tempered distribution (Appendix A.5.1).
Finally

$$(-1)^\theta D^{2\alpha} G_\alpha = G_0. \tag{1.48}$$

Proof. Since G_α has bounded support it follows from Appendix A.6.2 and Lemma 57 that

$$\widehat{G_\alpha} = \frac{1}{w_\alpha} \in C_{BP}^\infty, \tag{1.49}$$

and by Theorem 15

$$\frac{\xi^{2\alpha}}{w_\alpha(\xi)} \phi(\xi) \in S_{\emptyset,2\theta}, \quad \phi \in S. \tag{1.50}$$

$\boxed{\textbf{Case } 2\theta > d}$ Thus if ξ is the *action* variable

$$\left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta} w}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta} w_\alpha}, \xi^{2\alpha}\phi\right] = \left[\frac{1}{|\cdot|^{2\theta}}, \frac{\xi^{2\alpha}}{w_\alpha}\phi\right].$$

But Lemma 55 tells us that $T_\theta \in C_{BP}^{(2\theta-d-1)}$ and $\widehat{T_\theta} = \frac{c_\theta}{|\cdot|^{2\theta}}$ on $S_{\emptyset,2\theta}$. Hence

$$\left[\frac{1}{|\cdot|^{2\theta}}, \frac{\xi^{2\alpha}}{w_\alpha}\phi\right] = \left[\frac{1}{c_\theta}\widehat{T_\theta}, \frac{\xi^{2\alpha}}{w_\alpha}\phi\right] = \left[\frac{1}{c_\theta}\frac{\xi^{2\alpha}}{w_\alpha}\widehat{T_\theta}, \phi\right] = \left[\frac{1}{c_\theta}\xi^{2\alpha}\widehat{G_\alpha}\ \widehat{T_\theta}, \phi\right],$$

and, since G_α has bounded support, by part 4 of Appendix 245

$$\left[\frac{1}{c_\theta}\xi^{2\alpha}\widehat{G_\alpha}\ \widehat{T_\theta}, \phi\right] = \left[\frac{1}{c_\theta}\xi^{2\alpha}\widehat{G_\alpha * T_\theta}, \phi\right] = \left[\frac{(-1)^\theta}{c_\theta}\left(D^{2\alpha}G_\alpha * T_\theta\right)^\wedge, \phi\right],$$

so that $\widehat{G_\theta} = \frac{(-1)^\theta}{c_\theta}\left(D^{2\alpha}G_\alpha * T_\theta\right)^\wedge$ and $G_\theta = \frac{(-1)^\theta}{c_\theta}D^{2\alpha}G_\alpha * T_\theta$.

$\boxed{\textbf{Case } 2\theta < d \textbf{ and } d \textbf{ is odd}}$ If ξ is the distribution action variable then since $\frac{1}{w_\alpha} \in C_{BP}^\infty$

$$\left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta}\, w}, \phi\right] = \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)} w}, |\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi\right] = \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)} w_\alpha}, \xi^{2\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi\right]$$
$$= \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \frac{\xi^{2\alpha}}{w_\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi\right].$$

By Theorem 15, $\xi^{2\alpha} \in S_{\emptyset,2\theta}$, $|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)} \in S_{\emptyset,d+1-2\theta}$, $\xi^{2\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)} \in S_{\emptyset,d+1}$ and $\xi^{2\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi \in S_{\emptyset,d+1}$. Thus

$$\frac{\xi^{2\alpha}}{w_\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi \in S_{\emptyset,d+1}. \tag{1.51}$$

Also by Lemma 55,

$$\widehat{T_{\frac{d+1}{2}}} = c_{\frac{d+1}{2}}\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}} \quad on\ S_{\emptyset,d+1}, \tag{1.52}$$

so that

$$\left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \frac{\xi^{2\alpha}}{w_\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi\right] = \left[\frac{1}{c_{\frac{d+1}{2}}}\widehat{T_{\frac{d+1}{2}}}, \frac{\xi^{2\alpha}}{w_\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\phi\right]$$
$$= \left[\frac{1}{c_{\frac{d+1}{2}}}\xi^{2\alpha}|\cdot|^{2\left(\frac{d+1}{2}-\theta\right)}\frac{1}{w_\alpha}\widehat{T_{\frac{d+1}{2}}}, \phi\right]$$
$$= \left[\frac{1}{c_{\frac{d+1}{2}}}\xi^{2\alpha}|\cdot|^{d+1-2\theta}\widehat{G_\alpha}\ \widehat{T_{\frac{d+1}{2}}}, \phi\right]$$
$$= \left[\frac{1}{c_{\frac{d+1}{2}}}\xi^{2\alpha}|\cdot|^{d+1-2\theta}\left(G_\alpha * T_{\frac{d+1}{2}}\right)^\wedge, \phi\right]$$
$$= \left[\frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}}\left(|D|^{d+1-2\theta}D^{2\alpha}G_\alpha * T_{\frac{d+1}{2}}\right)^\wedge, \phi\right],$$

and hence

$$G_\theta = \frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}}|D|^{d+1-2\theta}D^{2\alpha}G_\alpha * T_{\frac{d+1}{2}}.$$

$\boxed{\textbf{Case } 2\theta \le d \textbf{ and } d \textbf{ is even}}$ This case can be proved in a very similar manner to the previous case. The last three cases combine to prove 1.46. Equation 1.47 follows directly from part 5 of Appendix 245.

Since $\widehat{G_\alpha} = \frac{1}{w_\alpha} \in L^1$ we have $x^{2\alpha}\widehat{G_\alpha} = \frac{1}{w} = \widehat{G_0}$, and hence $(-1)^\theta D^{2\alpha}G_\alpha = G_0 \in S' \cap \mathcal{E}'$. ■

Remark 59 *The reciprocal of an extended B-spline weight function is L^1 so the basis function G_0 is continuous. Substituting 1.48 i.e. $G_0 = (-1)^\theta D^{2\alpha}G_\alpha$ into 1.46 we obtain the equations 1.43 for the extended B-spline basis functions.*

We now will prove a more general result which **does not assume that G_α has bounded support**. Instead we assume that $G_\alpha \in L^1$ and that

$\int G_\alpha(y)\, T_m(x-y)\, dy$ defines a regular tempered distribution. To prove our theorem we will use the following result in which we use the function notation - a notational necessity - for the action of distributions:

Lemma 60 *(Theorem 2.7.5 Vladimirov [21]) Suppose $f, g \in \mathcal{D}'$ and g has bounded support ($g \in \mathcal{E}'$). Then the convolution $f * g$ exists and*

$$[f * g, \phi] = [f(x)\, g(y), \eta(y)\, \phi(x+y)], \quad \phi \in C_0^\infty, \tag{1.53}$$

where η is any test function equal to one in a neighborhood of $\operatorname{supp} g$.

Note: *the distribution on the right side of 1.53 is called a* ***tensor or direct product*** *and in the literature it is variously denoted $f \otimes g$, $f \times g$ or $f \cdot g$ in 'operator notation' and when 'function notation with action variables' is used we can write $f(x)\, g(y)$ or $f(x) \cdot g(y)$, for example. The tensor product is defined for $f, g \in \mathcal{D}'(\mathbb{R}^d)$ by*

$$[f \otimes g, \phi] = [f(x), [g(y), \phi(x,y)]], \quad \phi \in C_0^\infty(\mathbb{R}^d \times \mathbb{R}^d), \tag{1.54}$$

and has the properties you would expect e.g. commutivity, associativity, continuity. See for example Section 2.7 Vladimirov [21].

Theorem 61 *Suppose the weight function w has property W3.1 for order θ and κ. Then there exists a multi-index α such that $|\alpha| = \theta$ and 1.8 is true. Let G_α be the basis function of order zero generated by the zero order weight function $w_\alpha(\xi) = \xi^\alpha w(\xi)$ and assume $G_\alpha \in L^1$. Suppose T_m is the thin-plate spline basis function of integer order $m > d/2$ given in Lemma 55.*

We shall show that if G_θ is the basis function $\left(\frac{1}{|\cdot|^{2\theta} w}\right)^\vee$ of order θ generated by w, and H_θ is defined by the convolutions

$$H_\theta = \begin{cases} \frac{(-1)^\theta}{c_\theta} D^{2\alpha} G_\alpha * T_\theta, & \theta > d/2, \\ \frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} |D|^{d+1-2\theta} D^{2\alpha} G_\alpha * T_{\frac{d+1}{2}}, & \theta < d/2,\ d \text{ odd}, \\ \frac{(-1)^{\frac{d+2}{2}}}{c_{\frac{d+2}{2}}} |D|^{d+2-2\theta} D^{2\alpha} G_\alpha * T_{\frac{d+2}{2}}, & \theta \le d/2,\ d \text{ even}, \end{cases} \tag{1.55}$$

where

$$(G_\alpha * T_m)(x) = (2\pi)^{-d/2} \int G_\alpha(y)\, T_m(x-y)\, dy, \quad m > d/2, \tag{1.56}$$

then $G_\theta = H_\theta$ as tempered distributions whenever the convolution integral defines a regular tempered distribution.

Proof. In this proof we will refer to Lemma 57 for the properties of the zero order weight and basis functions used in this theorem. Since G_α no longer has bounded support we will use a partition of unity $\{\psi_\varepsilon, 1-\psi_\varepsilon\}$: choose $\psi \in C_0^\infty$ so that $\operatorname{supp}\psi \subset B(0;2)$, $0 \le \psi \le 1$ and $\psi = 1$ on $B(0;1)$. Now define $\psi_\varepsilon(\xi) = \psi(\varepsilon\xi)$ for $\varepsilon > 0$.

$\boxed{\textbf{Case } 2\theta > d}$ Since $\psi_\varepsilon G_\alpha$ has bounded support, $\widehat{\psi_\varepsilon G_\alpha} \in C_{BP}^\infty$ and if ξ is the *action* variable

$$\begin{aligned} \left[\widehat{G_\theta}, \phi\right] = \left[\frac{1}{|\cdot|^{2\theta} w}, \phi\right] &= \left[\frac{1}{|\cdot|^{2\theta} w_\alpha}, \xi^{2\alpha}\phi\right] \\ &= \left[\frac{\widehat{G_\alpha}}{|\cdot|^{2\theta}}, \xi^{2\alpha}\phi\right] \\ &= \left[\frac{\widehat{\psi_\varepsilon G_\alpha}}{|\cdot|^{2\theta}}, \xi^{2\alpha}\phi\right] + \left[\frac{((1-\psi_\varepsilon) G_\alpha)^\wedge}{|\cdot|^{2\theta}}, \xi^{2\alpha}\phi\right] \\ &= \left[\frac{1}{|\cdot|^{2\theta}}, \widehat{\psi_\varepsilon G_\alpha}\xi^{2\alpha}\phi\right] + \left[\frac{\xi^{2\alpha}((1-\psi_\varepsilon) G_\alpha)^\wedge}{|\cdot|^{2\theta}}, \phi\right]. \end{aligned}$$

But Lemma 55 tells us that $\widehat{T_\theta} = \frac{c_\theta}{|\cdot|^{2\theta}}$ on $S_{\emptyset,2\theta}$, and since $\widehat{\psi_\varepsilon G_\alpha}\xi^{2\alpha}\phi \in S_{\emptyset,2\theta}$ when $\phi \in S$

$$\left[\frac{1}{|\cdot|^{2\theta}}, \widehat{\psi_\varepsilon G_\alpha}\xi^{2\alpha}\phi\right] = \left[\frac{1}{c_\theta}\widehat{T_\theta}, \widehat{\psi_\varepsilon G_\alpha}\xi^{2\alpha}\phi\right] = \left[\frac{1}{c_\theta}\xi^{2\alpha}\widehat{\psi_\varepsilon G_\alpha}\ \widehat{T_\theta}, \phi\right]$$
$$= \left[\frac{(-1)^\theta}{c_\theta}\left(D^{2\alpha}\left(\psi_\varepsilon G_\alpha\right) * T_\theta\right)^\wedge, \phi\right],$$

so that

$$\widehat{G_\theta} = \frac{(-1)^\theta}{c_\theta}\left(D^{2\alpha}\left(\psi_\varepsilon G_\alpha\right) * T_\theta\right)^\wedge + \frac{\xi^{2\alpha}\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge}{|\cdot|^{2\theta}}. \tag{1.57}$$

We now want to show that the second term on the right converges to zero as a tempered distribution. In fact

$$\left|\left[\xi^{2\alpha}\frac{\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge}{|\cdot|^{2\theta}}, \phi\right]\right| = \left|\left[\xi^{2\alpha}\frac{\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge}{|\cdot|^{2\theta}}, \phi\right]\right| = \left|\int \xi^{2\alpha}\frac{\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge}{|\cdot|^{2\theta}}\phi\right|$$
$$\leq \int |\cdot|^{2\theta}\frac{\left|\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge\right|}{|\cdot|^{2\theta}}|\phi|$$
$$= \int \left|\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge\right||\phi|$$
$$\leq \left(\int \left|\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge\right|\right)\|\phi\|_\infty,$$

but since $G_\alpha \in L^1$ we have $\left|\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge\right| \leq \int_{|\cdot|\geq\frac{1}{\varepsilon}}|G_\alpha| \to 0$ as $\varepsilon \to 0^+$. Thus as a tempered distributions

$$\xi^{2\alpha}\frac{\left(\left(1-\psi_\varepsilon\right)G_\alpha\right)^\wedge}{|\cdot|^{2\theta}}\ as\ \varepsilon \to 0^+ \tag{1.58}$$

and so 1.57 implies that $\frac{(-1)^\theta}{c_\theta}\left(D^{2\alpha}\left(\psi_\varepsilon G_\alpha\right) * T_\theta\right)^\wedge$ converges to $\widehat{G_\theta}$ in the sense of tempered distribution i.e.

$$\frac{(-1)^\theta}{c_\theta}D^{2\alpha}\left(\psi_\varepsilon G_\alpha\right) * T_\theta \to G_\theta\ as\ \varepsilon \to 0^+. \tag{1.59}$$

Since $\psi_\varepsilon G_\alpha$ has bounded support we can use equation 1.53 of Lemma 60 to express the convolution in terms of the direct product. Thus

$$(2\pi)^{\frac{d}{2}}\left[\left(\psi_\varepsilon G_\alpha\right) * T_\theta, \phi\right] = \left[\left(\psi_\varepsilon G_\alpha\right)(x)\, T_\theta(y), \eta_\varepsilon(x)\,\phi(x+y)\right].$$

for any $\eta_\varepsilon \in C_0^\infty$ equal to one in a neighborhood of $\operatorname{supp}\left(\psi_\varepsilon G_\alpha\right)$. Now

$$\begin{aligned}
\left[\left(\psi_\varepsilon G_\alpha\right)(x)\, T_\theta(y), \eta_\varepsilon(x)\,\phi(x+y)\right] &= \int\int \psi_\varepsilon(x)\, G_\alpha(x)\, T_\theta(y)\,\phi(x+y)\, dydx \\
&= \int\int \psi_\varepsilon(x'-y')\, G_\alpha(x'-y')\, T_\theta(y')\,\phi(x')\, dy'dx' \\
&= \int\int \psi_\varepsilon(x'-y')\, G_\alpha(x'-y')\, T_\theta(y')\, dy'\phi(x')\, dx' \\
&= \int\int G_\alpha(x'-y')\, T_\theta(y')\, dy'\phi(x')\, dx' - \\
&\quad \int\limits_{\mathbb{R}^d}\ \int\limits_{|x'-y'|\geq 1/\varepsilon}\left(1-\psi_\varepsilon(x'-y')\right)G_\alpha(x'-y')\, T_\theta(y')\, dy'\phi(x')\, dx'.
\end{aligned}$$

Since we have assumed that $\int G_\alpha(x'-y')\, T_\theta(y')\, dy'$ defines a regular tempered distribution the first integral of the last equation exists and hence the second integral also exists and so converges to zero as $\varepsilon \to 0$ since the region of integration lies outside the ball $B\left(0; \frac{1}{\sqrt{d}\varepsilon}\right)$. Thus $\left(\psi_\varepsilon G_\alpha\right) * T_\theta$ converges to $(2\pi)^{-\frac{d}{2}}\int G_\alpha(x'-y')\, T_\theta(y')\, dy'$ as $\varepsilon \to 0$. But by 1.59, $H_\theta = \frac{(-1)^\theta}{c_\theta}D^{2\alpha}\left(\psi_\varepsilon G_\alpha\right) * T_\theta \to G_\theta$ as $\varepsilon \to 0^+$ and we can conclude that

$G_\theta = (2\pi)^{-\frac{d}{2}} \frac{(-1)^\theta}{c_\theta} D^{2\alpha} \int G_\alpha (x' - y') T_\theta (y') dy'$.

$\boxed{\textbf{Case } 2\theta < d \text{ and } d \text{ is odd}}$ If ξ is the *action* variable

$$\begin{aligned}
\left[\widehat{G_\theta}, \phi\right] &= \left[\frac{1}{|\cdot|^{2\theta} w}, \phi\right] = \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)} w}, |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)} w_\alpha}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{\widehat{G_\alpha}}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{\widehat{\psi_\varepsilon G_\alpha}}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] + \left[\frac{\left(\left(1-\psi_\varepsilon\right) G_\alpha\right)^\wedge}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{\widehat{\psi_\varepsilon G_\alpha}}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] + \left[\frac{\xi^{2\alpha}}{|\cdot|^{2\theta}} \left(\left(1-\psi_\varepsilon\right) G_\alpha\right)^\wedge, \phi\right],
\end{aligned} \tag{1.60}$$

so that by 1.51 and 1.52

$$\begin{aligned}
\left[\frac{\widehat{\psi_\varepsilon G_\alpha}}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] &= \left[\frac{1}{|\cdot|^{2\left(\frac{d+1}{2}\right)}}, \widehat{\psi_\varepsilon G_\alpha} \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{1}{c_{\frac{d+1}{2}}} \widehat{T_{\frac{d+1}{2}}}, \widehat{\psi_\varepsilon G_\alpha} \xi^{2\alpha} |\cdot|^{d+1-2\theta} \phi\right] \\
&= \left[\frac{1}{c_{\frac{d+1}{2}}} \xi^{2\alpha} |\cdot|^{d+1-2\theta} \widehat{\psi_\varepsilon G_\alpha}\ \widehat{T_{\frac{d+1}{2}}}, \phi\right] \\
&= \left[\frac{1}{c_{\frac{d+1}{2}}} \xi^{2\alpha} |\cdot|^{d+1-2\theta} \left(\psi_\varepsilon G_\alpha * T_{\frac{d+1}{2}}\right)^\wedge, \phi\right] \\
&= \left[\frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} \left(D^{2\alpha} |D|^{d+1-2\theta} \left(\psi_\varepsilon G_\alpha\right) * T_{\frac{d+1}{2}}\right)^\wedge, \phi\right].
\end{aligned}$$

and 1.60 now becomes

$$\left[\widehat{G_\theta}, \phi\right] = \left[\frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} \left(D^{2\alpha} |D|^{d+1-2\theta} \left(\psi_\varepsilon G_\alpha\right) * T_{\frac{d+1}{2}}\right)^\wedge, \phi\right] + \left[\frac{\xi^{2\alpha}}{|\cdot|^{2\theta}} \left(\left(1-\psi_\varepsilon\right) G_\alpha\right)^\wedge, \phi\right],$$

and thus

$$\widehat{G_\theta} = \frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} \left(D^{2\alpha} |D|^{d+1-2\theta} \left(\psi_\varepsilon G_\alpha\right) * T_{\frac{d+1}{2}}\right)^\wedge + \frac{\xi^{2\alpha}}{|\cdot|^{2\theta}} \left(\left(1-\psi_\varepsilon\right) G_\alpha\right)^\wedge.$$

Comparing this equation with equation 1.57 of the previous case we see that the subsequent calculations are also valid here and we can conclude that 1.58 holds i.e. $\xi^{2\alpha} \frac{\widehat{\left(1-\psi_\varepsilon\right) G_\alpha}}{|\cdot|^{2\theta}}$ converges to zero as a tempered distribution. Thus as tempered distributions

$$\frac{(-1)^{\frac{d+1}{2}}}{c_{\frac{d+1}{2}}} D^{2\alpha} |D|^{d+1-2\theta} \left(\psi_\varepsilon G_\alpha\right) * T_{\frac{d+1}{2}} \to G_\theta \ as\ \varepsilon \to 0^+.$$

Finally, the calculations of the previous part following 1.59 are valid for this case when $\theta = (d+1)/2$ and these complete the proof of this case.

$\boxed{\textbf{Case } 2\theta < d \text{ and } d \text{ is even}}$ This case can be proved in a very similar manner to the previous case. ■

Remark 62 *The condition $G_\alpha \in L^1$ of the last result can be used to deduce information about w_α and w. In fact, Corollary 3.7 of Petersen [16] implies that*

$$\frac{1}{w_\alpha(\xi)} = (2\pi)^{-\frac{d}{2}} \int e^{i\xi x} G_\alpha(x)\, dx \tag{1.61}$$

holds a.e. and that $\frac{1}{w_\alpha}$ can be modified on a set of measure zero so that it is a continuous, bounded function which converges to zero at infinity and is such that 1.61 holds everywhere. Thus

$$\frac{1}{w_\alpha(\xi)} \leq (2\pi)^{-\frac{d}{2}} \|G_\alpha\|_1 \tag{1.62}$$

and so $\lim_{|x|\to\infty} w_\alpha(x) = \infty$, $w_\alpha \in C^{(0)}(\mathbb{R}^d)$ and w_α is always positive which implies that w_α has property W1 for the set $\mathcal{A} = \{\}$.

Further, $\frac{1}{w(\xi)} = \frac{\xi^{2\alpha}}{w_\alpha(\xi)}$ so that by 1.62 $w(\xi) \geq \frac{(2\pi)^{\frac{d}{2}}}{\|G_\alpha\|_1}\frac{1}{\xi^\alpha}$ which implies w is always positive and can only have discontinuities on the axes $\xi_k = 0$. Indeed, w is discontinuous on $\xi_k = 0$ iff $\alpha_k = 0$.

The last theorem required that the convolution integral of 1.56 be a regular tempered distribution and that $G_\alpha \in L^1$. We complete this subsection by providing a single condition 1.63 for which those requirements are satisfied.

Corollary 63 *Suppose G_α and T_m are as given in Theorem 61 and suppose G_α also satisfies the inequality*

$$|G_\alpha(y)| \leq c(1+|y|)^{-s}, \tag{1.63}$$

for some $s > 2m$ and constant c. Then $G_\alpha \in L^1$ and

$$(2\pi)^{-d/2} \int G_\alpha(y)\, T_m(x-y)\, dy, \quad m > d/2, \tag{1.64}$$

is absolutely convergent for all x and defines a continuous function of polynomial increase. It is also a regular tempered distribution (as defined in Appendix A.5.1).

Proof. Since G_α satisfies inequality 1.63 and $m > d/2$ it follows that $s > d$ and so $G_\alpha \in L^1$. Further

$$\begin{aligned}
\left|\int G_\alpha(y)\, T_m(x-y)\, dy\right| &= \left|\int G_\alpha(x-y)\, T_m(y)\, dy\right| \\
&\leq \int |G_\alpha(x-y)\, T_m(y)|\, dy \\
&\leq c\int \frac{|T_m(y)|}{(1+|x-y|)^s} dy \\
&\leq c\int \frac{(1+|x|)^s}{(1+|y|)^s} |T_m(y)|\, dy \\
&= c(1+|x|)^s \int \frac{|T_m(y)|}{(1+|y|)^s} dy \\
&= c(1+|x|)^s \int_{|y|\leq 1} \frac{|T_m(y)|}{(1+|y|)^s} dy + c(1+|x|)^s \int_{|y|\geq 1} \frac{|T_m(y)|}{(1+|y|)^s} dy.
\end{aligned} \tag{1.65}$$

Since T_m is continuous the first integral of the last term exists. From 1.39

$$T_m(y) = \begin{cases} (-1)^{m-(d-2)/2} |y|^{2m-d} \log|y|, & d \text{ even}, \\ (-1)^{m-(d-1)/2} |y|^{2m-d}, & d \text{ odd}. \end{cases}$$

Thus if d is odd then

$$\int_{|y|\geq 1} \frac{|T_m(y)|}{(1+|y|)^s} dy \leq \int\limits_{|y|\geq 1} \frac{|y|^{2m-d}}{(1+|y|)^s} dy \leq \int\limits_{|y|\geq 1} |y|^{(2m-s)-d}\, dy,$$

which exists since $s > 2m$. If d is even and $s = 2m + 2\varepsilon$ then

$$\begin{aligned}
\int\limits_{|y|\geq 1} (1+|y|)^{-s} |T_m(y)|\, dy &\leq \int\limits_{|y|\geq 1} (1+|y|)^{-s} |y|^{2m-d} \log|y|\, dy \leq \int\limits_{|y|\geq 1} |y|^{2m-s-d} \log|y|\, dy \\
&= \int\limits_{|y|\geq 1} |y|^{-d-2\varepsilon} \log|y|\, dy \\
&= \int\limits_{|y|\geq 1} \frac{1}{|y|^{d+\varepsilon}} \frac{\log|y|}{|y|^{\varepsilon}} dy \\
&\leq \sup_{r\geq 1} \frac{\log r}{r^{\varepsilon}} \int\limits_{|y|\geq 1} \frac{1}{|y|^{d+\varepsilon}} \\
&< \infty.
\end{aligned}$$

We now know that if $s > 2m$ then for a constant c' independent of x

$$\int |G_\alpha(y)\, T_m(x-y)|\, dy \leq c' (1+|x|)^s, \quad x \in \mathbb{R}^d, \tag{1.66}$$

and so 1.64 is absolutely convergent and with polynomial growth at infinity.

To prove that $\int G_\alpha(y)\, T_m(x-y)\, dy$ is a continuous function of x we use the Lebesgue-dominated convergence theorem. First note that $\int G_\alpha(y)\, T_m(x-y)\, dy = \int G_\alpha(x-y)\, T_m(y)\, dy$ and that in the proof of 1.66 it was shown that

$$|G_\alpha(x-y)\, T_m(y)| \leq c\,(1+|x|)^s \frac{|T_m(y)|}{(1+|y|)^s},$$

and that $\frac{T_m(y)}{(1+|y|)^s} \in L^1$. Thus if $x_k \to x$ is any sequence in the ball $B(x;1)$ then

$$|G_\alpha(x_k-y)\, T_m(y)| \leq c\,(1+|x_k-x|+|x|)^s \frac{|T_m(y)|}{(1+|y|)^s} \leq c\,(2+|x|)^s \frac{|T_m(y)|}{(1+|y|)^s},$$

and by the Lebesgue-dominated convergence theorem

$$\lim_{x_k\to x} \int G_\alpha(y)\, T_m(x_k-y)\, dy = \lim_{x_k\to x} \int G_\alpha(x_k-y)\, T_m(y)\, dy = \int G_\alpha(x-y)\, T_m(y)\, dy,$$

and so proving continuity. A function $f \in L^1_{loc}$ for which $\int |f(x)|\,(1+|x|)^{-\lambda}\, dx$ exists for some $\lambda \geq 0$ is called a regular tempered distribution. Clearly a continuous function of polynomial growth at infinity is a regular tempered distribution. ■

1.6.6 *Positive definite and conditionally positive definite basis distributions*

In his work, Duchon did not mention the concept of conditional positive definite basis functions, but used as his starting point a weight function. Subsequent to Duchon's work various other classes of functions were used as basis functions in numerical work and these were studied from the point of view of (conditional) positive definiteness. This second major strand to the development of the theory of basis function interpolation has involved the use of the theory surrounding conditionally positive definite functions to develop the correct setting for a variational approach. Work on this approach includes papers by Schoenberg [19], Micchelli [15], Madych and Nelson [12], [13], Wu and Schaback [25] and in a series of papers by Schaback which the reader can find in the survey [18].

In their work Light and Wayne [11] returned to the approach of Duchon and the concept of (conditional) positive definiteness again plays a peripheral role. In the work of these authors it is only after their basis distributions have been defined, and their properties elucidated, that it is shown that they are positive definite or conditionally positive definite. They showed in Theorem 4.3 of [11] that the (integer) order of the basis distribution is the same as the order of the conditional positive definiteness.

Following Chapter 4 of Light and Wayne [11], the objective of this section is to show that the basis distributions G of order θ are conditionally positive definite distributions of order θ. The definition of a conditionally positive definite tempered distribution involves the concept of a homogeneous polynomial over the complex numbers $\mathbb{C}$.

Definition 64 ***Homogeneous polynomial over*** $\mathbb{C}$
A homogeneous polynomial p, of degree θ over $\mathbb{C}$, has the form

$$p(x) = \sum_{|\beta|=\theta} a_\beta x^\beta, \quad x \in \mathbb{R}^d,\ a_\beta \in \mathbb{C}.$$

Following Gelfand and Vilenkin [6] and Light and Wayne [11] we define a conditionally strictly positive definite tempered distribution as follows:

Definition 65 ***Conditionally strictly positive definite tempered distributions***
A distribution $F \in S'$ is said to be conditionally strictly positive definite of order $\theta \geq 1$ if the inequality $\left[p\bar{p}\widehat{F}, \psi\bar{\psi}\right] > 0$, holds for all $\psi \in S$, $\psi \neq 0$ and all homogeneous polynomials $p \neq 0$, of order θ.

Theorem 66 *Assume w is a weight function with property W2. Now suppose $G \in S'$ is a basis distribution of order θ generated by the weight function w.*
Then we show that G is a conditionally positive definite tempered distribution of order θ.

Proof. Suppose $p = \sum\limits_{|\beta|<\theta} a_\beta x^\beta$ is a homogeneous polynomial of order θ and $\psi \in S$, $\psi \neq 0$.

Then $\left[p\bar{p}\widehat{G}, \psi\bar{\psi}\right] = \left[\widehat{G}, p\bar{p}\psi\bar{\psi}\right]$ and

$$p\bar{p}\psi\bar{\psi} = \left(\sum_{|\alpha|=\theta} a_\alpha x^\alpha\right)\overline{\left(\sum_{|\beta|=\theta} a_\beta x^\beta\right)}\psi\bar{\psi} = \sum_{|\alpha|=\theta}\sum_{|\beta|=\theta} a_\alpha \overline{a_\beta} x^{\alpha+\beta}\psi\bar{\psi}.$$

Now $|\alpha+\beta| = 2\theta$ implies $x^{\alpha+\beta}\psi\bar{\psi} \in S_{\emptyset,2\theta}$ and so $p\bar{p}\psi\bar{\psi} \in S_{\emptyset,2\theta}$.
But G has order θ so, $\left[\widehat{G}, \phi\right] = \int \frac{\phi}{w|\cdot|^{2\theta}}$ for all $\phi \in S_{\emptyset,2\theta}$. Thus for all $\psi \in S$, $p\bar{p}\psi\bar{\psi} \in S_{\emptyset,2\theta}$ and

$$\left[p\bar{p}\widehat{G}, \psi\bar{\psi}\right] = \left[\widehat{G}, p\bar{p}\psi\bar{\psi}\right] = \int \frac{\left(p\bar{p}\psi\bar{\psi}\right)}{w|\cdot|^{2\theta}} = \int \frac{|p\psi|^2}{w|\cdot|^{2\theta}}.$$

Clearly $\int \frac{|p\psi|^2}{w|\cdot|^{2\theta}} = 0$ implies that $p\psi = 0$ a.e. Does this imply $p = 0$ or $\psi = 0$? Certainly, in one dimension p will have a finite number of zeros, and so $\psi = 0$. But what about in higher dimensions?

This result will be proved by induction on the number of dimensions. Suppose the result is true for d dimensions, and that $p\psi = 0$ in $d+1$ dimensions for some $p \neq 0$. Thus $p(x', x_d)\psi(x', x_d) = 0$ for all $x = (x', x_d)$. For fixed x_d, $p(x', x_d)$ is a polynomial in $d-1$ dimensions and $\psi(x', x_d) \in S$ in $d-1$ dimensions. Thus, for each x_d, $\psi(x', x_d) = 0$ for all x' and so $\psi = 0$. ■

1.7 Weight property generalizations

In this section we finish the document by briefly presenting two generalizations which include both properties W3.1 and W3.2.

Theorem 67 *Suppose $\theta \geq 1$ is an integer, $\kappa \geq 0$ and $r_3' \geq 0$. Suppose $\mu : \mathbb{R}^d \to \mathbb{R}$ and $0 < \mu(x) \leq |x|^{2\theta}$ a.e. on $|x| \geq r_3'$. Then if*

$$\int\limits_{|x|\geq r_3'} \frac{|x|^{2t}}{w(x)\mu(x)} dx < \infty, \quad 0 \leq t \leq \kappa, \tag{1.67}$$

we have:

1. *w has property W3.2 for θ, κ and r_3, where $r_3 = r_3'$ if $r_3' > 0$ and $r_3 > 0$ if $r_3' = 0$.*
2. *If $r_3' = 0$ then $1/w \in L^1_{loc}$ i.e. w has property W2.1.*
3. *If $\mu(x) = x^\alpha$, $|\alpha| = \theta$ and $r_3' = 0$ we have property W3.1.*
4. *If $\mu(x) = |x|^{2\theta}$ and $r_3' > 0$ we have property W3.2 for $r_3 = r_3'$.*

Proof. Part 1 The application of the inequality: $0 < \mu(x) \leq |x|^{2\theta}$ a.e. on $|x| \geq r_3$, to 1.67 implies the inequalities of part 1 of Theorem 7 and so w has property W3.2.

Part 2 If K is compact then $K \subset B(0;r)$ for some $r > 0$,and so

$$\int\limits_K \frac{1}{w} \leq \int\limits_{|\cdot| \leq r} \frac{1}{w} = \int\limits_{|\cdot| \leq r} \frac{\mu}{w\mu} \leq \int\limits_{|\cdot| \leq r} \frac{|\cdot|^{2\theta}}{w\mu} \leq r^{2\theta} \int\limits_{|\cdot| \leq r} \frac{1}{w\mu} < \infty,$$

by 1.67 with $r_3' = 0$ and $t = 0$.

Parts 3 and **4** follow from part 1 of Theorem 7. ■

Thus the smoothness results for X_w^θ and the basis distributions implied by property W3.2 also apply to property 1.67.

Now suppose $\theta \geq \phi \geq 1$ are integers, $\kappa \geq 0$ and $r_3' \geq 0$. We say w has weight function property WG if there exists a multi-index α such that $|\alpha| = \theta - \phi$ and

$$\int\limits_{|x| \geq r_3'} \frac{|x|^{2t}}{w(x)\,|x|^{2\phi}\,x^{2\alpha}} dx < \infty, \quad 0 \leq t \leq \kappa,$$

Clearly property WG generalizes properties W3.1 and W3.2 but the integral retains an interesting hybrid structure, part radial, part tensor product. With reference to the previous theorem, if we set $\mu(x) = |x|^{2\phi}\,x^{2\alpha}$ then $\mu(x) > 0$ a.e. when $|x| \geq r_3'$ and $\mu(x) = |x|^{2\phi}\,x^{2\alpha} = |x|^{2\phi}\left|x^{2\alpha}\right| \leq |x|^{2\phi}\,|x|^{2|\alpha|} = |x|^{2\theta}$. Thus $w(x) = |x|^{2\phi}\,x^{2\alpha}$ has property 1.67.

2

Proving a function is a basis function without using the spaces $S_{\emptyset,2n}$

2.1 Introduction

In this chapter we will prove some more properties of the basis functions studied in Chapter 1 where these objects were defined in terms of a weight function w and a positive integer order parameter θ. In Chapter 1 these weight function properties were used to define reproducing kernel semi-Hilbert spaces of continuous functions X_w^θ and continuous basis functions of order θ. In the Chapters 4, 5, 6 these semi-Hilbert spaces will be used to formulate and study several non-parametric, basis function interpolation and smoothing problems with the basis functions being used to represent the solutions to these problems.

In this chapter we are only concerned with the basis functions and we prove a result which will enable us to determine whether a given function is a basis function without recourse to the basis function definition 1.33 used by Light and Wayne in [11]. In practice their definition is difficult to use and involves a weight function and the bounded linear functionals on the subspace $S_{\emptyset,2\theta} = \{\phi \in S : D^\alpha \phi(0) = 0,\ |\alpha| < 2\theta\}$ of the test functions of the tempered distributions S (Appendix A.5). We give a simple test which can be applied to the tempered distribution Fourier transform of a continuous function.

These results are then applied to several classes of well-known radial basis functions, the choice here following Dyn [4]: the thin-plate splines, the shifted thin-plate splines, the multiquadric and inverse multiquadric functions and the Gaussian. These classes of basis functions are well known in the literature and details are given in Figure 2.2 below. In the last section I will illustrate the method using a non-radial example: the fundamental solutions of homogeneous elliptic differential operators of even order.

2.2 Theory

In this document we will prove that some of the important classes of functions used to define basis function interpolants and smoothers are basis functions in the Light sense i.e. generated by weight functions, without recourse to the awkward Definition 44 which uses the spaces $S_{\emptyset,2\theta}$ and $S'_{\emptyset,2\theta}$.

Our choice of basis functions is given in Table 2.2 below and follows Dyn [4]. These basis functions are well known in the literature. Theorem 69 is our main result and it will be applied to the various classes of radial and non-radial basis functions.

We will need the following basis distribution and weight function properties which were proved in the previous chapter:

Summary 68

1. *If a weight function w has property W2 then by part 2 of Appendix A.5.1, W2.1 and W2.2 imply that $1/w$ is a regular tempered distribution and thus $1/w \in L^1_{loc} \cap S'$.*

   *Now suppose the weight function w has properties W2.1 and **W3.2** for order θ and κ. Then:*

2. *w has property W2 (part 3 Theorem 10).*

3. *If G is a basis distribution of order θ then $|\cdot|^{2\theta}\widehat{G} = \frac{1}{w}$ as tempered distributions (part 2 Theorem 50).*

4. *$G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$, where $\lfloor \cdot \rfloor$ is the floor function (part 4 Theorem 50).*

The next theorem is the main result and will allow us to determine when a $C_{BP}^{(0)}$ function is a basis function in the Light sense by only studying its Fourier transform and so avoid the need to use the awkward Definition 44 of a basis function which uses $S_{\emptyset,2n}$ subspaces.

Theorem 69 *Suppose $H \in S'$ and $|\cdot|^{2\theta}\widehat{H} \in L^1_{loc}$. Hence for some $\mathcal{B} \subset \{0\}$ we can define the function $H_F \in L^1_{loc}(\mathcal{B})$ by $H_F = \widehat{H}$ on $\mathbb{R}^d \backslash \mathcal{B}$.*

Suppose further that for some closed set $\mathcal{A}$ of measure zero containing $\mathcal{B}$ we have $H_F \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ and $H_F(\xi) > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$. Now define the function w by

$$w(\xi) = \frac{1}{|\xi|^{2\theta} H_F(\xi)}, \quad \xi \in \mathbb{R}^d \backslash \mathcal{A}. \tag{2.1}$$

Then:

1. *w satisfies weight function property W1 w.r.t. the set $\mathcal{A}$.*

2. *Suppose w also has weight function properties W2.1 and **W3.2** for some order θ and κ. Then $H \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ and H is a basis function of order θ generated by w.*

Proof. Part 1 Clearly $\mathcal{A}$ is a closed set of measure zero and the properties of H imply that $w \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ and $w(\xi) > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$. Hence w has property W1 with w.r.t. the set $\mathcal{A}$.

Part 2. Since w has properties W2.1 and W3.2, part 2 of Summary 68 implies that w has property W2. Part 1 of Summary 68 then implies that $1/w \in L^1_{loc} \cap S'$.

By definition of H_F, $|\cdot|^{2\theta}\widehat{H} = |\cdot|^{2\theta} H_F$ as distributions on $\mathbb{R}^d \backslash \mathcal{A}$. By 2.1, $|\cdot|^{2\theta} H_F = 1/w$ a.e. so that $|\cdot|^{2\theta} H_F \in L^1_{loc}$ and since we have assumed that $|\cdot|^{2\theta}\widehat{H} \in L^1_{loc}$ it follows that $|\cdot|^{2\theta} H_F = |\cdot|^{2\theta}\widehat{H} = 1/w$ as distributions. But $|\cdot|^{2\theta}\widehat{H} \in S'$ and $1/w \in S'$ so $|\cdot|^{2\theta}\widehat{H} = 1/w$ as tempered distributions. If G is a basis distribution of order θ generated by w then by part 3 of Summary 68, $|\cdot|^{2\theta}\widehat{G} = 1/w$ and the basis (tempered) distribution definition implies $\widehat{G} = H_F$ on $\mathbb{R}^d \backslash 0$ as distributions. From the definition of H_F, $\widehat{H} = H_F$ on $\mathbb{R}^d \backslash 0$, and so $\operatorname{supp}\left(\widehat{G} - \widehat{H}\right) \subset \{0\}$, which implies $\widehat{G} - \widehat{H} = p$ where p is some polynomial. However, we have proved that $|\cdot|^{2\theta}\widehat{G} = 1/w$ and $|\cdot|^{2\theta}\widehat{H} = 1/w$ so, $0 = |\cdot|^{2\theta}\left(\widehat{G} - \widehat{H}\right) = |\cdot|^{2\theta}\widehat{p}$ i.e. $|D|^{2\theta} p = 0$. But $|D|^{2\theta}$ annihilates all members of $P_{2\theta}$ and no member of $P \backslash P_{2\theta}$ so we must have $p \in P_{2\theta}$. Thus $H \in G + P_{2\theta}$ and Definition 44 means that H is a basis distribution. Finally, by part 4 of Summary 68, $H \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$. ■

2.3 Examples: *radial basis functions* generated by weight functions in W3.2

Table 2.2 below gives a list of the radial functions derived from Dyn [4] and which are used as examples. Both types of splines have been generalized in the sense that the parameter s has been allowed to take non-integer values. Section 1 of Dyn [4] describes splines for which s has integer values. However, latter in Section 2 after Theorem 4, more general basis functions are introduced in equations 23 to 26. These are listed in Table 4.1, with a significant change. The change is to multiply by $(-1)^{\lceil s \rceil}$ or $(-1)^{s+1}$ so that the Fourier transform is positive.

Type	Radial basis function		
Surface/thin-plate spline (generalized s)	$(-1)^{\lceil s \rceil} r^{2s}$,	$s > 0,\ s \neq 1, 2, 3 \ldots .$	
	$(-1)^{s+1} r^{2s} \log r$,	$s = 1, 2, 3 \ldots .$	
Shifted surface /thin-plate spline (generalized s)	$(-1)^{\lceil s \rceil} \left(a^2 + r^2\right)^s$,	$a > 0,\ s > -d/2,$ $s \neq 1, 2, 3 \ldots .$	
	$\frac{(-1)^{s+1}}{2} \left(a^2 + r^2\right)^s \log \left(a^2 + r^2\right)$	$a > 0,$ $s = 1, 2, 3, \ldots .$	(2.2)
Multiquadric	$-\left(a^2 + r^2\right)^{1/2}$,	$a > 0,\ d > 1.$	
Inverse multi-quadric	$\left(a^2 + r^2\right)^{-1/2}$,	$a > 0.$	
Gaussian	$\exp\left(-\frac{r^2}{2}\right).$		

List of radial basis functions from Dyn [4].

2.3.1 *Thin-plate spline or surface spline functions*

For arbitrary dimension d the continuous thin-plate spline (or surface spline) function H can be defined by

$$H(x) = \begin{cases} (-1)^{s+1} |x|^{2s} \log |x|, & s = 1, 2, 3, \ldots, \\ (-1)^{\lceil s \rceil} |x|^{2s}, & s > 0,\ s \neq 1, 2, 3 \ldots . \end{cases} \tag{2.3}$$

Because H is a regular tempered distribution, $\widehat{H} \in S'$. In fact, from equations 23 and 24 of Dyn [4] we have that, as distributions,

$$\widehat{H} = e(s) |\cdot|^{-2s-d} \quad on\ \mathbb{R}^d \backslash 0, \tag{2.4}$$

where $e \in C_{BP}^{\infty}$. Specifically

$$e(s) = \begin{cases} (-1)^{s+1} c'(2s), & s = 1, 2, 3, \ldots, \\ (-1)^{\lceil s \rceil} c(2s), & s > 0,\ s \neq 1, 2, 3 \ldots, \end{cases} \tag{2.5}$$

where c and c' are defined by

$$c(t) = \pi^{d/2} 2^{t+d} \Gamma\left(\frac{t+d}{2}\right) / \Gamma\left(-\frac{t}{2}\right), \quad c' = \frac{dc}{dt}. \tag{2.6}$$

The next theorem will require that $\widehat{H}(\xi) > 0$ when $\xi \neq 0$ so the following lemma will be required.

Lemma 70 $e(s) > 0$ *when* $s > 0$.

Proof. The proof is an application of the reflection formula

$$\frac{1}{\Gamma(-x)} = -\frac{x \sin \pi x}{\pi} \Gamma(x).$$

First suppose that $s > 0$ and $s \neq 1, 2, 3 \ldots$. Then from 2.5 and 2.6

$$\begin{aligned} e(s) = (-1)^{\lceil s \rceil} c(2s) &= (-1)^{\lceil s \rceil} \pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) / \Gamma(-s) \\ &= (-1)^{\lceil s \rceil} \pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s)\, s \frac{-\sin \pi s}{\pi} \\ &= \pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s)\, s \frac{(-1)^{1+\lceil s \rceil} \sin \pi s}{\pi} \\ &> 0. \end{aligned}$$

Next assume $s = 1, 2, 3, \ldots$. Then

$$\begin{aligned} e(s) = (-1)^{s+1} c'(2s) &= (-1)^{s+1} \frac{1}{2} \frac{d}{ds} c(2s) \\ &= \frac{(-1)^{s+1}}{2} \frac{d}{ds} \left(\pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s)\, s \frac{-\sin \pi s}{\pi} \right) \\ &= \frac{(-1)^{s+1}}{2} \pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s)\, s\, (-\cos \pi s) + \\ &\quad + \frac{(-1)^{s+1}}{2} \frac{d}{ds} \left(\pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s)\, s \right) \frac{-\sin \pi s}{\pi} \\ &= \pi^{d/2} 2^{2s+d} \Gamma\left(s + \frac{d}{2}\right) \Gamma(s) \frac{s}{2} \\ &> 0, \end{aligned}$$

as required. ■

Theorem 71 *Let H be the generalized thin-plate function defined by equation 2.3. With reference to Theorem 69, equation 2.4 implies $H_F(\xi) = e(s) |\xi|^{-2s-d} \in L^1_{loc}\left(\mathbb{R}^d \backslash \mathcal{B}\right)$ where $\mathcal{B} = \{0\}$. For integer order $\theta \geq 1$ let $\mathcal{A} = \{0\}$ and define the function w by 2.1 i.e.*

$$w(\xi) = \frac{1}{|\xi|^{2\theta} H_F(\xi)} = \frac{1}{e(s)} |\xi|^{-2\theta+2s+d}, \quad \xi \neq 0. \tag{2.7}$$

Then:

1. *w has weight function property W1 for the set $\mathcal{A}$.*
2. *w also has weight function properties W2.1 and W3.2 for θ and $\kappa \geq 0$ iff $0 \leq \kappa < s < \theta$.*
3. *If $0 \leq \kappa < s < \theta$ then H is a basis function of order θ generated by w.*
4. *If $t > 2s$ there exists a constant c_t such that $|H(x)| \leq c_t (1 + |x|)^t$ for all x.*

Proof. This theorem is an application of Theorem . From 2.4 $\widehat{H}(\xi) = e(s) |\xi|^{-2s-d}$ so the condition $|\cdot|^{2\theta} \widehat{H} \in L^1_{loc}$ is satisfied iff $2\theta - 2s - d > -d$ i.e.

$$s < \theta. \tag{2.8}$$

Lemma 70 implies $H_F \in C^{(0)}\left(\mathbb{R}^d \backslash \mathcal{A}\right)$ and $H_F(\xi) > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$ so that if we define the function w by 2.7, part 1 of Theorem 69 implies w has **Property W1** w.r.t. the set $\mathcal{A}$.

Property W2.1 requires that $1/w \in L^1_{loc}$. But

$$\frac{1}{w(\xi)} = e(s) |\xi|^{2\theta-2s-d}, \tag{2.9}$$

so that $1/w \in L^1_{loc}$ iff $2\theta - 2s - d > -d$ i.e. iff $s < \theta$, which is already implied by 2.8.

Property W3.2 is true for order θ and κ if there exists $r_3 > 0$ such that $\int\limits_{|\cdot| \geq r_3} \frac{|\cdot|^{2\kappa}}{w |\cdot|^{2\theta}} < \infty$. But from 2.9, $\int\limits_{|\cdot| \geq r_3} \frac{|\cdot|^{2\kappa}}{w |\cdot|^{2\theta}} = \frac{1}{e(s)} \int\limits_{|\cdot| \geq r_3} \frac{1}{|\cdot|^{2s-2\kappa+d}}$ and this exists iff $2s - 2\kappa + d > d$ i.e. iff

$$\kappa < s. \tag{2.10}$$

Thus w has properties W1, W2.1 and W3.2 for some θ and κ if and only if $\theta > s$ and $0 \leq \kappa < s < \theta$. Now by part 2 of Theorem 69, H is a basis function of order θ generated by w.

From 2.3

$$(1 + |x|)^{-t} H(x) = \begin{cases} (-1)^{s+1} \frac{|x|^{2s} \log|x|}{(1+|x|)^t}, & s = 1, 2, 3, \ldots, \\ (-1)^{\lceil s \rceil} \frac{|x|^{2s}}{(1+|x|)^t}, & s > 0,\ s \neq 1, 2, 3 \ldots, \end{cases}$$

so that $t > 2s$ implies $(1 + |x|)^{-t} H(x) \to 0$ as $|x| \to \infty$. Thus $(1 + |x|)^{-t} |H(x)|$ is bounded. ■

2.3.2 Shifted thin-plate spline or shifted surface spline functions

For arbitrary dimension d the shifted thin-plate spline functions H are defined for $a > 0$ by

$$H(x) = \begin{cases} \frac{(-1)^{s+1}}{2}\left(a^2+|x|^2\right)^s \log\left(a^2+|x|^2\right), & s = 1,2,3,\ldots, \\ (-1)^{\lceil s\rceil}\left(a^2+|x|^2\right)^s, & s > -d/2,\ s \neq 1,2,3\ldots. \end{cases} \tag{2.11}$$

Now $H \in S'$ and from equations 25, 26 and 27 of Dyn [4], as distributions,

$$\widehat{H}(\xi) = \widetilde{e}(s)\,\widetilde{K}_{s+d/2}(a|\xi|)\,|\xi|^{-2s-d} \quad on\ \mathbb{R}^d\backslash 0,\ s > -d/2, \tag{2.12}$$

where

$$\widetilde{e}(s) = \begin{cases} (-1)^{s+1}\,\widetilde{c}'(2s), & s = 1,2,3,\ldots, \\ (-1)^{\lceil s\rceil}\,\widetilde{c}(2s), & s > -d/2,\ s \neq 1,2,3\ldots, \end{cases}$$

with

$$\widetilde{c}(t) = (2\pi)^{d/2}\,2^{\frac{t+2}{2}}/\Gamma(-t/2), \quad \widetilde{c}'(t) = \frac{d\widetilde{c}(t)}{dt}, \quad t > 0,$$

and

$$\widetilde{K}_\lambda(t) = t^\lambda K_\lambda(t), \quad t,\ \lambda\ real,$$

where K_λ is the modified Bessel function of the second kind of order λ. $\widetilde{K}_\lambda$ has the following properties

$$\widetilde{K}_\lambda \in C^{(0)}(\mathbb{R}); \quad \widetilde{K}_\lambda(t) > 0, \quad t \geq 0; \quad \lim_{t\to\infty}\widetilde{K}_\lambda(t) = 0\ exponentially. \tag{2.13}$$

See for example Abramowitz and Stegun [1]. The next theorem will require that $\widehat{H}(\xi) > 0$ when $\xi \neq 0$, so we will need the following lemma.

Lemma 72 $\widetilde{e}(s) > 0, \quad \xi \in \mathbb{R}^d,\ s > -d/2.$

Proof. The proof is very similar to the proof that $e(s) > 0$ when $s > 0$ (Lemma 70). It is again an application of the reflection formula 2.22.

First suppose that $s > -d/2$ and $s \neq 1,2,3\ldots$ Then

$$\begin{aligned} \widetilde{e}(s) = (-1)^{\lceil s\rceil}\,\widetilde{c}(2s) = (-1)^{\lceil s\rceil}(2\pi)^{d/2}\,2^{s+1}/\Gamma(-s) &= (-1)^{\lceil s\rceil}(2\pi)^{d/2}\,2^{s+1}\Gamma(s)\,s\frac{-\sin\pi s}{\pi} \\ &= (2\pi)^{d/2}\,2^{s+1}\Gamma(s)\,s\frac{(-1)^{1+\lceil s\rceil}\sin\pi s}{\pi} \\ &> 0. \end{aligned}$$

Next assume $s = 1,2,3,\ldots$ Then

$$\begin{aligned} e(s) = (-1)^{s+1}c'(2s) &= (-1)^{s+1}\frac{1}{2}\frac{d}{ds}c(2s) \\ &= \frac{(-1)^{s+1}}{2}\frac{d}{ds}\left(\pi^{d/2}2^{2s+d}\Gamma\left(s+\frac{d}{2}\right)\Gamma(s)\,s\frac{-\sin\pi s}{\pi}\right) \\ &= \frac{(-1)^{s+1}}{2}\pi^{d/2}2^{2s+d}\Gamma\left(s+\frac{d}{2}\right)\Gamma(s)\,s\,(-\cos\pi s)+ \\ &\quad + \frac{(-1)^{s+1}}{2}\frac{d}{ds}\left(\pi^{d/2}2^{2s+d}\Gamma\left(s+\frac{d}{2}\right)\Gamma(s)\,s\right)\frac{-\sin\pi s}{\pi} \\ &= \frac{1}{2}\pi^{d/2}2^{2s+d}\Gamma\left(s+\frac{d}{2}\right)\Gamma(s)\,s \\ &> 0, \end{aligned}$$

as required. ■

Theorem 73 *Let H be the generalized shifted thin-plate function defined by equation 2.11. With reference to Theorem 69, equation 2.12 together with 2.13 imply $H_F(\xi) = \widetilde{e}(s)\,\widetilde{K}_{s+d/2}(a|\xi|)\,|\xi|^{-2\theta+2s+d} \in L^1_{loc}\left(\mathbb{R}^d\backslash\mathcal{B}\right)$ where $\mathcal{B} = \{0\}$. For integer order $\theta \geq 1$ let $\mathcal{A} = \{0\}$ and define the function w by 2.1 i.e.*

$$w(\xi) = \frac{1}{|\xi|^{2\theta}H_F(\xi)} = \frac{1}{\widetilde{e}(s)\,\widetilde{K}_{s+d/2}(a|\xi|)}\,|\xi|^{-2\theta+2s+d}, \quad s > -d/2. \tag{2.14}$$

Then:

1. *w has weight function property W1 w.r.t. the set $\mathcal{A}$.*
2. *w also has weight function properties W2.1 and W3.2 for θ and all $\kappa \geq 0$ iff $-d/2 < s < \theta$.*
3. *If $-d/2 < s < \theta$ then H is a basis function of order θ generated by w.*
4. *If $t > 2s$ there exists a constant c_t such that $|H(x)| \leq c_t (1+|x|)^t$ for all x.*

Proof. This theorem is an application of Theorem 69. From 2.12, $\widehat{H}(\xi) = \widetilde{e}(s)\widetilde{K}_{s+d/2}(a|\xi|)|\xi|^{-2s-d}$, $\xi \neq 0$, $s > -d/2$, so the condition $|\cdot|^{2\theta}\widehat{H} \in L^1_{loc}$ is satisfied iff $2\theta - 2s - d > -d$ and $s > -d/2$ i.e.

$$-d/2 < s < \theta. \tag{2.15}$$

Lemma 72 and 2.13 imply $H_F \in C^{(0)}\left(\mathbb{R}^d\backslash\mathcal{A}\right)$ and $H_F(\xi) > 0$ on $\mathbb{R}^d\backslash\mathcal{A}$ so that if we define the function w by 2.14, part 1 of Theorem 69 implies w has **Property W1** w.r.t. the set $\mathcal{A}$.

Property W2.1 requires that $1/w \in L^1_{loc}$. But

$$\frac{1}{w(\xi)} = \widetilde{e}(s)\widetilde{K}_{s+d/2}(a|\xi|)|\xi|^{2\theta-2s-d}, \quad s > -d/2 \tag{2.16}$$

and 2.13 implies $\widetilde{K}_\lambda \in C^{(0)}(\mathbb{R})$ and $\widetilde{K}_\lambda(t) > 0$ for $t \geq 0$, so that $1/w \in L^1_{loc}$ iff $|\xi|^{2\theta-2s-d} \in L^1_{loc}$ iff $2\theta - 2s - d > -d$ i.e. $s < \theta$, which is again constraint 2.15.

Property W3.2 is true for order θ and κ if there exists $r_3 > 0$ such that $\int_{|\cdot|\geq r_3} \frac{|\cdot|^{2\kappa}}{w|\cdot|^{2\theta}} < \infty$. But by 2.16, when $r_3 > 0$

$$\int_{|\cdot|\geq r_3} \frac{|\cdot|^{2\kappa}}{w|\cdot|^{2\theta}} = \widetilde{e}(s)\int_{|\xi|\geq r_3} \widetilde{K}_{s+d/2}(a|\xi|)|\xi|^{2\kappa-2s-d}\,d\xi,$$

and since by 2.13, $\widetilde{K}_{s+d/2}(a|\xi|)$ is continuous and $\lim\limits_{|\xi|\to\infty} \widetilde{e}(s)\widetilde{K}_{s+d/2}(a|\xi|) \to 0$ exponentially, this integral always exists. Thus w has properties W1, W2.1 and W3.2 for some θ and any $\kappa \geq 0$ if and only 2.15 is satisfied. Now by part 2 of Theorem 69, H is a basis function of order θ generated by w.

From 2.11

$$(1+|x|)^{-t} H(x) = \begin{cases} (-1)^{s+1} \frac{\left(a^2+|\xi|^2\right)^s \log\left(a^2+|\xi|^2\right)}{(1+|x|)^t}, & s = 1,2,3,\ldots, \\ (-1)^{\lceil s \rceil} \frac{\left(a^2+|\xi|^2\right)^s}{(1+|x|)^t}, & s > 0,\ s \neq 1,2,3\ldots, \end{cases}$$

so that $t > 2s$ implies $(1+|x|)^{-t} H(x) \to 0$ as $|x| \to \infty$. Thus $(1+|x|)^{-t}|H(x)|$ is bounded. ■

2.3.3 *Multiquadric and inverse multiquadric basis functions*

In arbitrary dimension the multiquadric function is defined by:

$$H(\xi) = \left(a^2 + |\xi|^2\right)^{1/2}, \quad a > 0.$$

In dimension $d \geq 2$ the inverse multiquadric function is defined by:

$$H(\xi) = \left(a^2 + |\xi|^2\right)^{-1/2}, \quad a > 0.$$

Note that these are clearly specific cases of the shifted surface splines discussed in the previous subsection. These functions were introduced by Hardy [7] for geophysical applications. See also the review paper by Hardy [8].

2.3.4 *The Gaussian*

The Gaussian function is

$$H(x) = \exp\left(-|x|^2\right), \quad x \in \mathbb{R}^d. \tag{2.17}$$

and has Fourier transform

$$\widehat{H}(\xi) = \frac{\sqrt{\pi}}{2}\exp\left(-|\xi|^2/4\right), \quad \xi \in \mathbb{R}^d. \tag{2.18}$$

Theorem 74 *Let H be the Gaussian function defined by equation 2.17. With reference to Theorem 69, equation 2.18 implies*

$$H_F(\xi) = \widehat{H}(\xi) \in L^1_{loc}(\mathbb{R}^d), \tag{2.19}$$

so $\mathcal{B} = \{\}$. For integer order $\theta \geq 1$ let $\mathcal{A} = \{0\}$ and define the function w by 2.1 i.e.

$$w(\xi) = \frac{1}{|\xi|^{2\theta} H_F(\xi)} = \frac{2}{\sqrt{\pi}} |\xi|^{-2\theta} \exp\left(|\xi|^2/4\right), \quad \xi \notin \mathcal{A}. \tag{2.20}$$

Then:

1. *w has weight function property W1 w.r.t. the set $\mathcal{A}$.*
2. *w also has weight function properties W2.1 and W3.2 for all θ and $\kappa \geq 0$.*
3. *H is a basis function of order θ generated by w.*
4. *For any real t there exists a constant c_t such that $|H(x)| \leq c_t (1 + |x|)^t$ for all x.*

Proof. This theorem is an application of Theorem 69. From 2.18 the condition $|\cdot|^{2\theta} \widehat{H} \in L^1_{loc}$ is always satisfied. Equations 2.18 and 2.19 imply $H_F \in C^{(0)}(\mathbb{R}^d \backslash \mathcal{A})$ and $H_F(\xi) > 0$ on $\mathbb{R}^d \backslash \mathcal{A}$ so that if we define the function w by 2.7, part 1 of Theorem 69 means that w has **Property W1** w.r.t. the set $\mathcal{A} = \{0\}$.

Property W2.1 requires that $1/w \in L^1_{loc}$. But

$$\frac{1}{w(\xi)} = \frac{2}{\sqrt{\pi}} |\xi|^{2\theta} \exp\left(-|\xi|^2/4\right), \tag{2.21}$$

so this is clearly true.

Property W3.2 is true for given order θ and κ if there exists $r_3 > 0$ such that $\int\limits_{|\cdot| \geq r_3} \frac{|\cdot|^{2\kappa}}{w |\cdot|^{2\theta}} < \infty$, but this is obviously true from 2.21.

Thus w has properties W1, W2.1 and W3.2 for any order θ and κ and by part 2 of Theorem 69, H is a basis function of order θ generated by w.

Finally, it is clear that $(1 + |x|)^{-t} |H(x)|$ is bounded for any real t.. ■

2.4 Examples: *non-radial basis functions* generated by weight functions in W3.2

2.4.1 *Fundamental solutions of homogeneous elliptic differential operators of even order*

In Section 2 of Dyn [4] describes a large class of non-radial functions on $\mathbb{R}^d$ which are positive definite, namely the fundamental solutions of homogeneous elliptic differential operators of even order 2θ, where $2\theta > d$. In the next result we will use the weight function theory to prove that these fundamental solutions are scalar multiples of basis functions of order θ generated by a weight function.

Theorem 75 *Suppose p is a homogeneous polynomial of degree 2θ on $\mathbb{R}^d$, where $\theta \geq 1$ and $2\theta > d$. Further, suppose that $p(x) > 0$ if $x \neq 0$, and that $f \in S'$ is a fundamental solution of the differential operator $p(D)$. Then:*

1. *The function w defined by*

$$w(x) = p\left(\frac{x}{|x|}\right), \quad x \neq 0, \tag{2.22}$$

has property W1 of a weight function w.r.t. the set $\mathcal{A} = \{0\}$. It also has properties W2.1 and W3.2 for order θ and κ iff $0 \leq 2\kappa < 2\theta - d$.

2. *$H = (-1)^\theta (2\pi)^{d/2} f$ satisfies $|\cdot|^{2\theta} \widehat{H} = \frac{1}{w} \in L^1_{loc}$ as distributions.*
3. *$H \in C^{(2\theta-d-1)}_{BP}$ is a basis function of order θ generated by the weight function w.*

Proof. Part 1 Since the sphere $|x|=1$ is a compact set

$$c_1=\min\{p(x):|x|=1\},\quad c_2=\max\{p(x):|x|=1\},$$

both exist. Also, $c_1,c_2>0$ since $p(x)>0$ if $x\neq 0$. We conclude that

$$0<c_1\leq w(x)\leq c_2,\quad x\neq 0,$$

and clearly

$$w\in C_B^{(0)}\left(\mathbb{R}^d\backslash 0\right);\quad 0<\frac{1}{c_2}\leq\frac{1}{w(x)}\leq\frac{1}{c_1},\quad x\neq 0, \tag{2.23}$$

and

$$\frac{1}{p(x)}=\frac{1}{w(x)\,|x|^{2\theta}},\quad x\neq 0. \tag{2.24}$$

Thus w has property W1 of a weight function for the set $\mathcal{A}=\{0\}$, as well as property W2.1. Property W3.2 holds for order θ and κ iff there exists $r_3>0$ such that $\int_{|\cdot|\geq r_3}\frac{|\cdot|^{2\kappa}}{w\,|\cdot|^{2\theta}}<\infty$ which by 2.23 is true iff $2\theta-2\kappa>d$. Thus w satisfies properties W1, W2.1 and W3.2 for order θ and κ iff $0\leq 2\kappa<2\theta-d$.

Part 2 A tempered distribution f which satisfies the equation $p(D)f=\delta$ is said to be a fundamental solution of the differential operator $p(D)$. Taking the Fourier transform and noting that p is homogeneous of degree 2θ, $p(D)f=p(-i\xi)\widehat{f}=(-1)^{\theta}p\widehat{f}=\widehat{\delta}=(2\pi)^{-d/2}$, so that $(-1)^{\theta}p\widehat{f}=(2\pi)^{-d/2}$ i.e. $p\widehat{H}=1$ as tempered distributions.

But 2.23 implies $\frac{1}{w}\in L^1_{loc}$ and since $\frac{1}{w}=\frac{|\cdot|^{2\theta}}{w|\cdot|^{2\theta}}=\frac{|\cdot|^{2\theta}}{p}=|\cdot|^{2\theta}\widehat{H}$ a.e. it follows that $\frac{1}{w}=|\cdot|^{2\theta}\widehat{H}$ as distributions.

Part 3 Here we will make use of Theorem 69. Now $H\in S'$ and by part 2, $|\cdot|^{2\theta}\widehat{H}=\frac{1}{w}\in L^1_{loc}$ as tempered distributions so for $\mathcal{B}=\{0\}$ we can define $H_F\in L^1_{loc}(\mathcal{B})$ by $H_F=\widehat{H}$ on $\mathbb{R}^d\backslash\mathcal{B}$. From the proof of part 2, $p\widehat{H}=1$ as tempered distributions. Thus if we set $\mathcal{A}=\mathcal{B}$ it follows that $H_F=1/p$ on $\mathbb{R}^d\backslash\mathcal{A}$ and $H_F\in C^{(0)}\left(\mathbb{R}^d\backslash\mathcal{A}\right)$ and $H_F(\xi)>0$ on $\mathbb{R}^d\backslash\mathcal{A}$. Equation 2.24 implies the weight function w of part 1 coincides with that of 2.1 in Theorem 69, and so the results of part 1 allow us to use Theorem 69 to conclude that $H\in C_{BP}^{(\lfloor 2\kappa\rfloor)}$ and H is a basis function of order θ generated by w. But $0\leq 2\kappa<2\theta-d$ implies $\lfloor 2\kappa\rfloor\leq 2\theta-d-1$ so $H\in C_{BP}^{(2\theta-d-1)}$. ∎

3

More basis function and semi-Hilbert data space theory

3.1 Introduction

This chapter is based on the theory developed in Chapter 1 which extended the work of Light and Wayne [11] The results given here are also closely related to the work of Madych and Nelson [13] as is indicated in the remarks to several of our theorems. However, Madych and Nelson deal with the *reciprocal* of our weight function so that *their* X_w^θ space is formally $X_{1/w}^{-\theta} = \left\{ u \in S' : \int \frac{|\widehat{u}|^2}{w|\cdot|^{2\theta}} < \infty \right\}$ so that in order to calculate their data space they need to construct a mapping from $X_{1/w}^{-\theta}$ to X_w^θ Section 3. In a future document I will discuss the space $X_{1/w}^{-\theta}$ and construct isometric isomorphisms from X_w^θ to $X_{1/w}^{-\theta}$ which show that $X_{1/w}^{-\theta}$ is isomorphic to the space of bounded linear functionals on X_w^θ i.e. $\left(X_w^\theta\right)' = X_{1/w}^{-\theta}$.

In Chapter 1 smoothness and growth properties were derived for basis functions and smoothness, growth and completeness properties were exhibited for the data spaces X_w^θ. It was shown that every weight function generates a $C^{(\lfloor 2\kappa \rfloor)}$ basis function where $\lfloor \cdot \rfloor$ denotes the floor function. However, the rate of growth of basis functions near infinity was only determined for a subclass of the weight functions. This subclass generates bounded basis functions with bounded derivatives. As for the functions in X_w^θ, it was shown that $X_w^\theta \subset C_{BP}^{(\lfloor \kappa \rfloor)}$ for all weight functions but no bounds on growth rates near infinity were obtained.

The goal of this chapter is to derive 'modified' inverse-Fourier transform formulas for basis functions and the data functions X_w^θ and to use these formulas to obtain bounds for the rates of increase of these functions and their derivatives near infinity. This will be done by proving a general inverse-Fourier transform formula for a subspace of the distributions and then applying it to both the basis functions and the data functions. From these formulas we will be able to show that all basis functions are either bounded or have a rate of increase of at most $|\cdot|^{\lfloor 2\theta \rfloor}$ near infinity and that all data functions have a rate of increase of at most $|\cdot|^{\lfloor \kappa \rfloor}$ near infinity.

We will also show that:

1. There always exists a conjugate-even (complex-valued) basis function.
2. If the weight function is even then there exists an even, real valued basis function.
3. If the weight function is radial then there exists a radial basis function.

The key operators used in this document are the projections $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ which are given by

$$\left(\mathcal{P}_{\emptyset,n} u\right)(x) = \rho(x) \sum_{|\alpha|<n} \frac{x^\alpha}{\alpha!} D^\alpha u(0), \quad \mathcal{Q}_{\emptyset,n} = I - \mathcal{P}_{\emptyset,n},$$

where $\rho \in S$, $\rho(0) = 1$, $D^\beta \rho(0) = 0$ for $1 \le |\beta| < n$. Here S is the space of C^∞ test functions for the tempered distributions - see Appendix A.5. Thus they are based on the Taylor series expansion about the origin.

Section by section in brief:

In **Section** 3.2 the null spaces and ranges of $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ are derived.

In **Section** 3.3 the distribution adjoints of $\mathcal{P}_{\emptyset,n}^*$ and $\mathcal{Q}_{\emptyset,n}^*$ are calculated. These are projections and we calculate their null spaces and ranges.

In **Section** 3.4 we prove some upper bounds for the function $\left|D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right|$.

In **Section** 3.5 an inverse Fourier transform result is derived for a class of distributions which will be applied in the next two sections.

In **Section** 3.6 we prove an inverse Fourier transform formula for the basis functions. This is formula is used to derive the rates of increase of the basis functions near infinity and to derive the properties of a basis function given various properties of the weight function.

In **Section** 3.7 an inverse Fourier transform theorem is proved for functions in X_w^θ and then used to estimate their rates of increase near infinity.

3.2 The operators $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n} = I - \mathcal{P}_{\emptyset,n}$

The space $S_{1,n}$ is required to define the operators $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ which are central to this document:

Definition 76 ***The spaces*** $S_{1,n}$

$$S_{1,n} = \begin{cases} S, & n = 0, \\ \left\{\phi \in S : \phi(0) = 1;\ D^\beta \phi(0) = 0, \quad 1 \leq |\beta| < n\right\}, & n = 1, 2, 3, \ldots \end{cases}$$

Definition 77 ***The space*** ρP_n ***and the operators*** $\mathcal{P}_{\emptyset,n}$, $\mathcal{Q}_{\emptyset,n}$.
Suppose $\rho \in S_{1,n}$. *Then for integer* $n \geq 0$ *we define:*

1. *The space* $\rho P_n = \{\rho p : p \in P_n\}$.

2. *The mappings* $\mathcal{P}_{\emptyset,n} : C^{(n)} \to C^{(n)}$ *and* $\mathcal{Q}_{\emptyset,n} = I - \mathcal{P}_{\emptyset,n} : C^{(n)} \to C^{(n)}$ *are given by*

$$\mathcal{P}_{\emptyset,n} u(x) = \rho(x) \sum_{|\alpha| < n} \frac{x^\alpha}{\alpha!} D^\alpha u(0).$$

$$\mathcal{Q}_{\emptyset,n} u(x) = u(x) - \rho(x) \sum_{|\alpha| < n} \frac{x^\alpha}{\alpha!} D^\alpha u(0).$$

The next theorem proves some important properties of the operators $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ as well as demonstrating relationships between S and $S_{\emptyset,n}$.

Theorem 78 *Suppose* $\rho \in S_{1,n}$. *Then:*

1. $\mathcal{P}_{\emptyset,n}$ *and* $\mathcal{Q}_{\emptyset,n}$ *are continuous linear mappings from* S *into* S.

2. $\mathcal{P}_{\emptyset,n}$ *and* $\mathcal{Q}_{\emptyset,n}$ *are projections into* S *which satisfy:*

$$\begin{array}{lll} \mathcal{P}_{\emptyset,n} : S \to \rho P_n & \textit{is onto and} & \mathrm{null}\, \mathcal{P}_{\emptyset,n} = S_{\emptyset,n}, \\ \mathcal{Q}_{\emptyset,n} : S \to S_{\emptyset,n} & \textit{is onto and} & \mathrm{null}\, \mathcal{Q}_{\emptyset,n} = \rho P_n. \end{array}$$

3. $S = S_{\emptyset,n} \oplus \rho P_n$.

4. $S = \widehat{S_{\emptyset,n}} \oplus \widehat{\rho P_n} = \widehat{S_{\emptyset,n}} \oplus \{p(D)\widehat{\rho} : p \in P_n\}$.

Proof. Part 1 It needs to be shown that if $\phi_k \to 0$ in S then $\mathcal{P}_{\emptyset,n}\phi_k \to 0$ in S
i.e. $x^\alpha D^\beta\left(\mathcal{P}_{\emptyset,n}\phi_k\right)(x) \to 0$ for all x, α, β. Since

$$\mathcal{P}_{\emptyset,n}\phi_k(x) = \rho(x) \sum_{|\alpha| < n} \frac{x^\alpha}{\alpha!} \left(D^\alpha \phi_k\right)(0) = \rho(x) \sum_{|\alpha| < n} \frac{\left[D^\alpha \delta, \phi_k\right]}{\alpha!} x^\alpha,$$

we have using Leibniz's rule A.1 in the Appendix.

$$\begin{aligned}
D^{\beta}\left(\mathcal{P}_{\emptyset,n}\phi_k\right)(x) &= D^{\beta}\left(\rho(x)\sum_{|\alpha|<n}\frac{[D^{\alpha}\delta,\phi_k]}{\alpha!}x^{\alpha}\right) \\
&= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\left(D^{\beta-\gamma}\rho\right)(x)\,D^{\gamma}\left(\sum_{|\alpha|<n}\frac{[D^{\alpha}\delta,\phi_k]}{\alpha!}x^{\alpha}\right) \\
&= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\left(D^{\beta-\gamma}\rho\right)(x)\left(\sum_{|\alpha|<n}\frac{[D^{\alpha}\delta,\phi_k]}{\alpha!}D^{\gamma}x^{\alpha}\right) \\
&= \sum_{\gamma\leq\beta}\binom{\beta}{\gamma}\left(D^{\beta-\gamma}\rho\right)(x)\left(\sum_{\substack{|\alpha|<n\\ \alpha\leq\gamma}}\frac{[D^{\alpha}\delta,\phi_k]}{\alpha!}D^{\gamma}x^{\alpha}\right).
\end{aligned}$$

If $\phi_k \to 0$ in S then $[D^{\alpha}\delta,\phi_k] \to 0$, since $D^{\alpha}\delta \in S'$. Hence $D^{\beta}\left(\mathcal{P}_{\emptyset,n}\phi_k\right)(x) \to 0$ for all x and so $x^{\alpha}D^{\beta}\left(\mathcal{P}_{\emptyset,n}\phi_k\right)(x) \to 0$ for all x, α, β.

Parts 2 and 3

$$\begin{aligned}
\mathcal{P}_{\emptyset,n}\mathcal{P}_{\emptyset,n}\phi = \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}\left(D^{\beta}\mathcal{P}_{\emptyset,n}\phi\right)(0) &= \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}D^{\beta}\left(\rho(x)\sum_{|\alpha|<n}\frac{x^{\alpha}}{\alpha!}D^{\alpha}\phi(0)\right)(0) \\
&= \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}\rho(x)\,D^{\beta}\left(\sum_{|\alpha|<n}\frac{x^{\alpha}}{\alpha!}D^{\alpha}\phi(0)\right)(0),
\end{aligned}$$

since $(D^{\alpha}\rho)(0) = 0$ for $1 \leq |\alpha| < n$. Further, $\rho(0) = 1$ implies

$$\begin{aligned}
\mathcal{P}_{\emptyset,n}\mathcal{P}_{\emptyset,n}\phi = \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}D^{\beta}\left(\sum_{|\alpha|<n}\frac{x^{\alpha}}{\alpha!}D^{\alpha}\phi(0)\right)(0) &= \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}\left(\frac{D^{\beta}\left(x^{\beta}\right)}{\beta!}D^{\beta}\phi(0)\right)(0) \\
&= \rho(x)\sum_{|\beta|<n}\frac{x^{\beta}}{\beta!}D^{\beta}\phi(0) \\
&= \mathcal{P}_{\emptyset,n}\phi(x),
\end{aligned}$$

and so $\mathcal{P}_{\emptyset,n}$ is a projection. Hence $\mathcal{Q}_{\emptyset,n}$ is also a projection, $S = \operatorname{null}\mathcal{Q}_{\emptyset,n} \oplus \operatorname{range}\mathcal{P}_{\emptyset,n}$ and clearly $\operatorname{range}\mathcal{P}_{\emptyset,n} = \rho P_n$.

Now to show that $\operatorname{null}\mathcal{P}_{\emptyset,n} = S_{\emptyset,n}$. Since $\rho(0) = 1$ there exists $r > 0$ such that $\rho(x) > 0$ for $|x| < r$. Hence if $\mathcal{P}_{\emptyset,n}\phi = 0$

$$\rho(x)\sum_{|\alpha|<n}\frac{x^{\alpha}}{\alpha!}D^{\alpha}\phi(0) = \phi(x), \quad |x| < r,$$

and so $\sum\limits_{|\alpha|<n}\frac{x^{\alpha}}{\alpha!}D^{\alpha}\phi(0) = 0$ for $|x| < r$. Thus $D^{\alpha}\phi(0) = 0$ if $|\alpha| < n$, and so $\phi \in S_{\emptyset,n}$. Conversely, $\phi \in S_{\emptyset,n}$ clearly implies $\mathcal{P}_{\emptyset,n}\phi = 0$ and since $\mathcal{P}_{\emptyset,n} = I - \mathcal{Q}_{\emptyset,n}$ we have the results

$$\begin{aligned}
\operatorname{range}\mathcal{Q}_{\emptyset,n} &= \operatorname{null}\mathcal{P}_{\emptyset,n} = S_{\emptyset,n}, \\
\operatorname{null}\mathcal{Q}_{\emptyset,n} &= \operatorname{range}\mathcal{P}_{\emptyset,n} = \rho P_n.
\end{aligned}$$

■

3.3 The tempered distribution adjoints of $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$

The projections $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ were studied in the previous section and here we will derive their tempered distribution adjoints $\mathcal{P}^{*}_{\emptyset,n}$ and $\mathcal{Q}^{*}_{\emptyset,n}$. These are projections into S' and several properties are proved which relate to their ranges and null spaces.

Definition 79 ***The distribution adjoints*** $\mathcal{P}^*_{\emptyset,n}$ ***and*** $\mathcal{Q}^*_{\emptyset,n}$

The adjoints of $\mathcal{P}_{\emptyset,n}, \mathcal{Q}_{\emptyset,n} : S \to S$ *are denoted by* $\mathcal{P}^*_{\emptyset,n}, \mathcal{Q}^*_{\emptyset,n}$ *and are defined for* $u \in S'$ *and* $\varphi \in S$ *by*

$$\begin{aligned} \left[\mathcal{Q}^*_{\emptyset,n} u, \varphi\right] &= \left[u, \mathcal{Q}_{\emptyset,n}\varphi\right], \\ \left[\mathcal{P}^*_{\emptyset,n} u, \varphi\right] &= \left[u, \mathcal{P}_{\emptyset,n}\varphi\right], \end{aligned}$$

so that $\mathcal{P}^*_{\emptyset,n} : S' \to S'$ *and* $\mathcal{Q}^*_{\emptyset,n} : S' \to S'$ *are continuous.*

Theorem 80 *The operators* $\mathcal{P}^*_{\emptyset,n}$ *and* $\mathcal{Q}^*_{\emptyset,n}$ *have the following properties:*

1. $\mathcal{P}^*_{\emptyset,n}$ *and* $\mathcal{Q}^*_{\emptyset,n}$ *are projections and* $\mathcal{P}^*_{\emptyset,n} + \mathcal{Q}^*_{\emptyset,n} = I$.

2. *If* $u \in S'$ *then* $\mathcal{P}^*_{\emptyset,n} u = p_u(-iD)\,\delta$ *where*

$$p_u(\xi) = \sum_{|\alpha|<n} \frac{b_{\alpha,u}}{\alpha!}\xi^\alpha \in P_n, \quad b_{\alpha,u} = \left[u, (-ix)^\alpha \rho\right].$$

3. $\operatorname{null} \mathcal{Q}^*_{\emptyset,n} = \operatorname{range} \mathcal{P}^*_{\emptyset,n} = \{p(D)\,\delta : p \in P_n\} = \widehat{P}_n$.

4. $\operatorname{range} \mathcal{Q}^*_{\emptyset,n} = \operatorname{null} \mathcal{P}^*_{\emptyset,n} = \{u \in S' : [u, x^\alpha \rho] = 0 \text{ when } |\alpha| < n\}$.

5. $S' = \widehat{P}_n \oplus \{u \in S' : [u, x^\alpha \rho] = 0 \text{ when } |\alpha| < n\}$.

6. $\varphi u = \varphi \mathcal{Q}^*_{\emptyset,n} u$, *where* $u \in S'$ *and* $\varphi \in C^\infty_{\emptyset,n} \cap C^{(0)}_{BP}$.

Proof. Part 1 is true since $\mathcal{P}_{\emptyset,n}$ and $\mathcal{Q}_{\emptyset,n}$ are projections.

Part 2 If $u \in S'$ and $\varphi \in S$ then

$$\begin{aligned} \left[u, \mathcal{P}_{\emptyset,n}\varphi\right] = \left[u, \rho \sum_{|\alpha|<n} \frac{x^\alpha}{\alpha!} D^\alpha \varphi(0)\right] &= \sum_{|\alpha|<n} \frac{[u, x^\alpha \rho]}{\alpha!} D^\alpha \varphi(0) \\ &= \sum_{|\alpha|<n} \frac{[u, x^\alpha \rho]}{\alpha!} \left[\delta, D^\alpha \varphi\right] \\ &= \sum_{|\alpha|<n} \frac{[u, x^\alpha \rho]}{\alpha!} (-1)^{|\alpha|} \left[D^\alpha \delta, \varphi\right] \\ &= \left[\left(\sum_{|\alpha|<n} \frac{[u, x^\alpha \rho]}{\alpha!} (-1)^{|\alpha|} D^\alpha\right) \delta, \varphi\right] \\ &= \left[\left(\sum_{|\alpha|<n} \frac{[u, (-ix)^\alpha \rho]}{\alpha!} (-iD)^\alpha\right) \delta, \varphi\right] \\ &= \left[\mathcal{P}^*_{\emptyset,n} u, \varphi\right], \end{aligned}$$

and thus

$$\mathcal{P}^*_{\emptyset,n} u = \left(\sum_{|\alpha|<n} \frac{b_{\alpha,u}}{\alpha!} (-iD)^\alpha\right) \delta \in S'.$$

Part 3 From the formulas of part 2 it is clear that $\mathcal{P}^*_{\emptyset,n} : S' \to \widehat{P}_n$ so that $\operatorname{range} \mathcal{P}^*_{\emptyset,n} \subset \widehat{P}_n$. We now show that $\operatorname{null} \mathcal{Q}^*_{\emptyset,n} = \widehat{P}_n$. If $u \in \widehat{P}_n$ then $u = p(D)\delta$ for some $p \in P_n$ and, since $\varphi \in S$ implies $\mathcal{Q}_{\emptyset,n}\varphi \in S_{\emptyset,n}$,

$$\begin{aligned} \left[\mathcal{Q}^*_{\emptyset,n} u, \varphi\right] = \left[u, \mathcal{Q}_{\emptyset,n}\varphi\right] = \left[p(D)\,\delta, \mathcal{Q}_{\emptyset,n}\varphi\right] &= \left[\delta, p(-D)\left(\mathcal{Q}_{\emptyset,n}\varphi\right)\right] \\ &= \left[p(-D)\left(\mathcal{Q}_{\emptyset,n}\varphi\right)\right](0) \\ &= 0, \end{aligned}$$

by the definition of $S_{\emptyset,n}$. This means $\mathcal{Q}^*_{\emptyset,n}u = 0$ and hence that $\widehat{P}_n \subset \operatorname{null}\mathcal{Q}^*_{\emptyset,n}$.

Part 4 From part 2, $\mathcal{P}^*_{\emptyset,n}u = p_u(D)\,\delta$ so $\mathcal{P}^*_{\emptyset,n}u = 0$ iff $\widehat{\mathcal{P}^*_{\emptyset,n}u} = 0$ iff $p_u = 0$ iff $[u, x^\alpha\rho] = 0$ when $|\alpha| < n$. Here x was an *action* variable - see Notation 1.

Part 5 Part 1 implies that $S' = \operatorname{null}\mathcal{Q}^*_{\emptyset,n} \oplus \operatorname{range}\mathcal{Q}^*_{\emptyset,n}$.

Part 6 Suppose $u \in S'$ and $\varphi \in C^\infty_{\emptyset,n} \cap C^{(0)}_{BP}$. Then $\psi \in S$ implies $\varphi\psi \in S_{\emptyset,n}$ and so

$$[\varphi u, \psi] = [u, \varphi\psi] = \left[u, \mathcal{Q}_{\emptyset,n}(\varphi\psi)\right] = \left[\mathcal{Q}^*_{\emptyset,n}u, \varphi\psi\right] = \left[\varphi\mathcal{Q}^*_{\emptyset,n}u, \psi\right].$$

■

3.4 Bounds for $\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)$

In Section 3.2 the operator $\mathcal{Q}_{\emptyset,n}$ was introduced and now we will derive some upper bounds for the function $\left|\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right|$, where $\mathcal{Q}_{\emptyset,n,\xi}$ acts on the variable ξ in $e^{i(x,\xi)}$. These estimates will be used in Section 3.6 to prove an explicit formula for a basis function in terms of its weight function, and used again in Section 3.7 to prove an inverse Fourier transform for a member of X^θ_w.

Theorem 81 *The function $\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)$ has the following properties:*

1. *There exists a constant $C_{n,\gamma}$, independent of x and ξ, such that*

$$\left|D^\gamma_x\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right| \le \begin{cases} C_{n,\gamma}\,|\xi|^{|\gamma|}\,(1+|x|)^{n-|\gamma|-1}, & |\gamma| < n, \\ |\xi|^{|\gamma|}, & |\gamma| \ge n. \end{cases}$$

   This is a useful estimate for large x and large ξ.

2. *Given $r > 0$ there exists a constant $C_{n,\gamma,r}$ such that for all x and $|\xi| \le r$*

$$\left|D^\gamma_x\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right| \le \begin{cases} C_{n,\gamma,r}\,|\xi|^{n}\,(1+|x|)^{n-|\gamma|}, & |\gamma| < n, \\ |\xi|^{|\gamma|}, & |\gamma| \ge n. \end{cases} \tag{3.1}$$

   This estimate is useful near $\xi = 0$.

Proof. We first derive a formula for $D^\gamma_x\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)$. For some $\rho \in S_{1,n}$

$$\begin{aligned}\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) &= e^{i(x,\xi)} - \rho(\xi)\sum_{|\alpha|<n}\frac{i^{|\alpha|}x^\alpha\xi^\alpha}{\alpha!} = e^{i(x,\xi)} - \rho(\xi)\sum_{k<n}\sum_{|\alpha|=k}\frac{i^{|\alpha|}x^\alpha\xi^\alpha}{\alpha!} \\ &= e^{i(x,\xi)} - \rho(\xi)\sum_{k<n}\frac{(ix,\xi)^k}{k!},\end{aligned}$$

so that

$$D^\gamma_x\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) = \begin{cases} (i\xi)^\gamma\,\mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right), & |\gamma| < n, \\ (i\xi)^\gamma, & |\gamma| \ge n. \end{cases} \tag{3.2}$$

Part 1 If $|\gamma| \ge n$ then the inequality is a very simple consequence of the formula for $D^\gamma_x\mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) = (i\xi)^\gamma\,e^{i(x,\xi)}$.

On the other hand, if $|\gamma| < n$ then from equations 3.2

$$\begin{aligned}
\left|D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right| &= \left|(i\xi)^\gamma \mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right)\right| \\
&\leq |\xi|^{|\gamma|}\left(1+\sum_{k<n-|\gamma|}\frac{|\rho(\xi)|\,|\xi|^k\,|x|^k}{k!}\right) \\
&\leq |\xi|^{|\gamma|}\left(1+\sum_{k<n-|\gamma|}\frac{|\rho(\xi)|\,|\xi|^k\,(1+|x|)^k}{k!}\right) \\
&\leq |\xi|^{|\gamma|}\left(1+\sum_{k<n-|\gamma|}\frac{|\rho(\xi)|\,|\xi|^k}{k!}\right)(1+|x|)^{n-|\gamma|-1} \\
&\leq \max_{j<n-|\gamma|}\left\||\cdot|^j\rho\right\|_\infty |\xi|^{|\gamma|}\left(1+\sum_{k<n-|\gamma|}\frac{1}{k!}\right)(1+|x|)^{n-|\gamma|-1} \\
&< 4\max_{j<n-|\gamma|}\left\||\cdot|^j\rho\right\|_\infty |\xi|^{|\gamma|}(1+|x|)^{n-|\gamma|-1} \\
&= C_{n,\gamma}|\xi|^{|\gamma|}(1+|x|)^{n-|\gamma|-1},
\end{aligned}$$

where $C_{n,\gamma} = 4\max\limits_{k<n-|\gamma|}\left\||\cdot|^k\rho\right\|_\infty$, as required.

Part 2 If $|\gamma| \geq n$ the inequality follows from part 1. There remains the case $|\gamma| < n$. Using the Taylor series expansion about zero (Appendix A.8) we define the function μ_m by $e^{it} = \sum\limits_{k<m}\frac{(it)^k}{k!} + (it)^m\mu_m(t)$ and note that $\|\mu_m\|_\infty \leq \frac{1}{m!}$. The following calculations

$$\begin{aligned}
D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) = (i\xi)^\gamma \mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right) &= (i\xi)^\gamma\left(e^{i(x,\xi)}-\rho(\xi)\sum_{k<n-|\gamma|}\frac{(ix,\xi)^k}{k!}\right) \\
&= (i\xi)^\gamma\left(e^{i(x,\xi)}-\rho(\xi)\left(e^{i(x,\xi)}-(ix,\xi)^{n-|\gamma|}\mu_{n-|\gamma|}((x,\xi))\right)\right) \\
&= (i\xi)^\gamma\left(e^{i(x,\xi)}(1-\rho(\xi))+\rho(\xi)(ix,\xi)^{n-|\gamma|}\mu_{n-|\gamma|}((x,\xi))\right),
\end{aligned}$$

now imply that when $|\xi| \leq r$

$$\begin{aligned}
\left|D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right| &\leq |\xi|^{|\gamma|}\left(|1-\rho(\xi)|+|\rho(\xi)|\,|\xi|^{n-|\gamma|}\,|x|^{n-|\gamma|}\right) \\
&\leq \left(|\xi|^{|\gamma|}\frac{|1-\rho(\xi)|}{|\xi|^n}+|\rho(\xi)|\right)|\xi|^n(1+|x|)^{n-|\gamma|} \\
&\leq \left(r^{|\gamma|}\frac{|1-\rho(\xi)|}{|\xi|^n}+\|\rho\|_\infty\right)|\xi|^n(1+|x|)^{n-|\gamma|}.
\end{aligned}$$

But $\rho \in S_{1,n}$ implies $1-\rho \in C_{\emptyset,n}^\infty$ and by Taylor's theorem (Appendix A.8) there exists a constant $k_{\rho,n} > 0$, independent of ξ, such that $|1-\rho(\xi)| \leq k_{\rho,n}|\xi|^n$ for all ξ. Thus

$$\left|D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\right| \leq C_{n,\gamma,r}|\xi|^n(1+|x|)^{n-|\gamma|},$$

where $C_{n,\gamma,r} = r^{|\gamma|}k_{\rho,n} + \|\rho\|_\infty$. ■

3.5 A useful inverse Fourier transform theorem

The next theorem will be used in Section 3.6 to derive an inverse Fourier transform formula for a basis function and in Section 3.7 to derive an inverse Fourier transform formula for any member of X_w^θ. This theorem will use the following lemma which shows how, when $\phi \in S$, the expressions $\mathcal{P}_{\emptyset,n}\phi$ and $\mathcal{Q}_{\emptyset,n}\phi$ can be expressed in terms of $\mathcal{P}_{\emptyset,n,x}\left(e^{i(x,\xi)}\right)$ and $\mathcal{Q}_{\emptyset,n,x}\left(e^{i(x,\xi)}\right)$.

Lemma 82 *If $\psi \in S$ then:*

1. $\left(\mathcal{P}_{\emptyset,n}\overset{\vee}{\psi}\right)(\xi) = (2\pi)^{-\frac{d}{2}} \int \mathcal{P}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx.$
2. $\left(\mathcal{Q}_{\emptyset,n}\overset{\vee}{\psi}\right)(\xi) = (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx.$

Proof. If $\psi \in S$ then

$$\begin{aligned}
\left(\mathcal{P}_{\emptyset,n}\overset{\vee}{\psi}\right)(\xi) = \rho(\xi)\sum_{|\alpha|<n}\frac{\xi^\alpha D^\alpha\overset{\vee}{\psi}(0)}{\alpha!} &= \rho(\xi)\sum_{|\alpha|<n}\frac{\xi^\alpha}{\alpha!}\left((ix)^\alpha\psi\right)^\vee(0) \\
&= (2\pi)^{-\frac{d}{2}}\rho(\xi)\sum_{|\alpha|<n}\frac{\xi^\alpha}{\alpha!}\int (ix)^\alpha\psi(x)\,dx \\
&= (2\pi)^{-\frac{d}{2}}\int\left(\sum_{|\alpha|<n}\rho(\xi)\frac{(ix)^\alpha\xi^\alpha}{\alpha!}\right)\psi(x)\,dx \\
&= (2\pi)^{-\frac{d}{2}}\int \mathcal{P}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx,
\end{aligned}$$

and hence

$$\begin{aligned}
\left(\mathcal{Q}_{\emptyset,n}\overset{\vee}{\psi}\right)(\xi) &= \overset{\vee}{\psi}(\xi) - \left(\mathcal{P}_{\emptyset,n}\overset{\vee}{\psi}\right)(\xi) \\
&= (2\pi)^{-\frac{d}{2}}\int e^{i\xi x}\psi(x)\,dx - (2\pi)^{-d/2}\int \mathcal{P}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx \\
&= (2\pi)^{-\frac{d}{2}}\int \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx \\
&= (2\pi)^{-\frac{d}{2}}\int \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right)\psi(x)\,dx.
\end{aligned}$$

■

To prove the next theorem we will require another lemma concerning differentiation under the integral sign.

Lemma 83 *(Prop 7.8.4 of Malliavin [14])Suppose $f : \mathbb{R}^{m+n} \to \mathbb{C}$ and we write $f(\xi, x)$ where $\xi \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$. Further suppose that:*

1. *For each ξ, $f(\xi,\cdot) \in C^{(k)}(\mathbb{R}^n)$.*
2. *For each x, $\int\left|D_\xi^\alpha f(\xi,x)\,d\xi\right| < \infty$ for $|\alpha| \le k$.*

Then we have

$$D_x^\alpha \int f(\xi,x)\,d\xi = \int D_x^\alpha f(\xi,x)\,d\xi, \text{ when } |\alpha| \le k,$$

and $\int f(\xi,\cdot)\,d\xi \in C^{(k)}(\mathbb{R}^n)$.

Theorem 84 *Suppose $f \in S'$ and $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d\backslash 0\right)$. Define the function $f_F : \mathbb{R}^d \to \mathbb{C}$ a.e. by $f_F = \widehat{f}$ on $\mathbb{R}^d\backslash 0$. Then:*

1. *If the action $\int f_F\phi$ on $\phi \in S_{\emptyset,n}$ defines a member of $S'_{\emptyset,n}$, and if $f_F = \widehat{f}$ on $S_{\emptyset,n}$, it follows that for all multi-indexes γ*

$$\left[\widehat{D^\gamma f},\psi\right] = \int (i\xi)^\gamma\left(\mathcal{Q}_{\emptyset,n}\psi\right)(\xi)\,f_F(\xi)\,d\xi + (2\pi)^{-\frac{d}{2}}\left[p_{\widehat{D^\gamma f}},\psi\right], \quad \psi \in S, \tag{3.3}$$

where for $u \in S'$, $p_u \in P_n$ is defined by

$$p_u(x) = \sum_{|\alpha|<n}\frac{b_{\alpha,u}}{\alpha!}x^\alpha, \quad b_{\alpha,u} = \left[u,(-i\xi)^\alpha\rho\right]. \tag{3.4}$$

Further, $\widehat{D^\gamma f} = (i\xi)^\gamma f_F$ on $S_{\emptyset,n}$, and $(i\xi)^\gamma f_F \in L^1_{loc}\left(\mathbb{R}^d\backslash 0\right)$.

2. *Now also assume that for a given multi-index γ there exists $s \geq 0$ and a constant $k > 0$ independent of ξ such that*
$$\int \left| D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)} \right) f_F(\xi) \right| d\xi \leq k \left(1 + |x|\right)^s. \tag{3.5}$$
Then $D^\gamma f \in C_{BP}^{(0)}$,
$$D^\gamma f(x) = (2\pi)^{-\frac{d}{2}} \int D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)} \right) f_F(\xi)\, d\xi + (2\pi)^{-\frac{d}{2}} D^\gamma p_{\widehat{f}}(x), \tag{3.6}$$
and
$$\int D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)} \right) f_F(\xi)\, d\xi = \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n-|\gamma|,\xi} \left(e^{i(x,\xi)} \right) (i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| < n, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| \geq n, \end{cases} \tag{3.7}$$
$$= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n-|\gamma|,\xi} \left(e^{i(x,\xi)} \right) (D^\gamma f)_F(\xi)\, d\xi, & |\gamma| < n, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (D^\gamma f)_F(\xi)\, d\xi, & |\gamma| \geq n, \end{cases} \tag{3.8}$$
where the function $(D^\gamma f)_F : \mathbb{R}^d \to \mathbb{C}$ is a.e. defined by $(D^\gamma f)_F = \widehat{D^\gamma f}$ on $\mathbb{R}^d \backslash 0$.

Proof. Part 1 We first prove equation 3.3. From part 2 of Theorem 78 we know that $\mathcal{Q}_{\emptyset,n} : S \to S_{\emptyset,n}$ and since the current theorem assumes that $f_F = \widehat{f}$ on $S_{\emptyset,n}$ and that $\int f_F \phi$ on $\phi \in S_{\emptyset,n}$ defines a member of $S'_{\emptyset,n}$, we have for $\psi \in S$ and the *action* variable ξ
$$\left[\widehat{D^\gamma f}, \psi \right] = \left[(i\xi)^\gamma \widehat{f}, \psi \right] = \left[(i\xi)^\gamma \widehat{f}, \mathcal{Q}_{\emptyset,n}\psi \right] + \left[(i\xi)^\gamma \widehat{f}, \mathcal{P}_{\emptyset,n}\psi \right].$$
But
$$\left[(i\xi)^\gamma \widehat{f}, \mathcal{Q}_{\emptyset,n}\psi \right] = \left[\widehat{f}, (i\xi)^\gamma \mathcal{Q}_{\emptyset,n}\psi \right] = \left[f_F, (i\xi)^\gamma \mathcal{Q}_{\emptyset,n}\psi \right] = \int (i\xi)^\gamma \mathcal{Q}_{\emptyset,n}\psi(\xi) f_F(\xi)\, d\xi,$$
and
$$\left[(i\xi)^\gamma \widehat{f}, \mathcal{P}_{\emptyset,n}\psi \right] = \left[\mathcal{P}_{\emptyset,n}^* \left((i\xi)^\gamma \widehat{f} \right), \psi \right] = \left[\mathcal{P}_{\emptyset,n}^* \left(\widehat{D^\gamma f} \right), \psi \right],$$
so that
$$\left[\widehat{D^\gamma f}, \psi \right] = \int (i\xi)^\gamma \mathcal{Q}_{\emptyset,n}\psi(\xi) f_F(\xi)\, d\xi + \left[\mathcal{P}_{\emptyset,n}^* \left(\widehat{D^\gamma f} \right), \psi \right]. \tag{3.9}$$
The next step is to simplify the second term on the right. From part 1 of Theorem 80
$$\mathcal{P}_{\emptyset,n}^* \widehat{f} = \left(\sum_{|\alpha|<n} \frac{b_{\alpha,\widehat{f}}}{\alpha!} (-iD)^\alpha \right) \delta, \quad b_{\alpha,\widehat{f}} = \left[\widehat{f}, (-i\xi)^\alpha \rho \right].$$
Hence
$$\left(\mathcal{P}_{\emptyset,n}^* \widehat{f} \right)^\vee (x) = (2\pi)^{-\frac{d}{2}} \sum_{|\alpha|<n} \frac{b_{\alpha,\widehat{f}}}{\alpha!} x^\alpha = (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}(x),$$
so that
$$\left[\mathcal{P}_{\emptyset,n}^* \widehat{f}, \psi \right] = \left[\left(\mathcal{P}_{\emptyset,n}^* \widehat{f} \right)^\vee, \widehat{\psi} \right] = (2\pi)^{-\frac{d}{2}} \left[p_{\widehat{f}}, \widehat{\psi} \right] = (2\pi)^{-\frac{d}{2}} \left[\widehat{p_{\widehat{f}}}, \psi \right],$$
and
$$\mathcal{P}_{\emptyset,n}^* \left(\widehat{D^\gamma f} \right) = (2\pi)^{-\frac{d}{2}} \widehat{p_{\widehat{D^\gamma f}}}, \tag{3.10}$$
so that 3.9 becomes 3.3. It remains to show that $\widehat{D^\gamma f} = (i\xi)^\gamma f_F$ on $S_{\emptyset,n}$. In fact, if $\phi \in S_{\emptyset,n}$ then, since $f_F = \widehat{f}$ on $S_{\emptyset,n}$, $\left[\widehat{D^\gamma f}, \phi \right] = \left[(i\xi)^\gamma \widehat{f}, \phi \right] = \left[\widehat{f}, (i\xi)^\gamma \phi \right] = \left[f_F, (i\xi)^\gamma \phi \right] = \left[(i\xi)^\gamma f_F, \phi \right]$ as required. Also, we are given that $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$ and $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ so $f_F \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ and $(i\xi)^\gamma f_F \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$.

Part 2 The next step is to prove equation 3.6. If $\psi \in S$ then since $\mathcal{Q}_{\emptyset,n} + \mathcal{P}_{\emptyset,n} = I$

$$\begin{aligned}[D^\gamma f, \psi] = (-1)^{|\gamma|} [f, D^\gamma \psi] = (-1)^{|\gamma|} \left[\widehat{f}, (D^\gamma \psi)^\vee\right] = (-1)^{|\gamma|} \left[\widehat{f}, \mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)\right] + \\ + (-1)^{|\gamma|} \left[\widehat{f}, \mathcal{P}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)\right].\end{aligned}$$

By part 2 of Theorem 78 $\mathcal{Q}_{\emptyset,n} : S \to S_{\emptyset,n}$ and since this theorem assumes that $f_F = \widehat{f}$ on $S_{\emptyset,n}$ and that $\int f_F \phi$ on $\phi \in S_{\emptyset,n}$ defines a member of $S'_{\emptyset,n}$, we have

$$\left[\widehat{f}, \mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)\right] = \left[f_F, \mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)\right] = \int \mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)(\xi)\, f_F(\xi)\, d\xi.$$

Also by 3.10 with $\gamma = 0$

$$\left[\widehat{f}, \mathcal{P}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)\right] = \left[\left(\mathcal{P}^*_{\emptyset,n} \widehat{f}\right)^\vee, D^\gamma \psi\right] = (2\pi)^{-\frac{d}{2}} \left[p_{\widehat{f}}, D^\gamma \psi\right] = (2\pi)^{-\frac{d}{2}} \left[(-D)^\gamma p_{\widehat{f}}, \psi\right],$$

so that now

$$[D^\gamma f, \psi] = (-1)^{|\gamma|} \int \mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)(\xi)\, f_F(\xi)\, d\xi + (2\pi)^{-\frac{d}{2}} \left[D^\gamma p_{\widehat{f}}, \psi\right].$$

From Lemma 82, for action variable x

$$\begin{aligned}\mathcal{Q}_{\emptyset,n} \left((D^\gamma \psi)^\vee\right)(\xi) &= (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) D^\gamma \psi(x)\, dx \\ &= (2\pi)^{-\frac{d}{2}} \left[\mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(\cdot,\xi)}\right), D^\gamma \psi\right] \\ &= (2\pi)^{-\frac{d}{2}} \left[(-D)^\gamma \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(\cdot,\xi)}\right), \psi\right] \\ &= (2\pi)^{-\frac{d}{2}} \int (-D)^\gamma_x \left(\mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right)\right) \psi(x)\, dx,\end{aligned}$$

so that

$$\begin{aligned}[D^\gamma f, \psi] &= (2\pi)^{-\frac{d}{2}} \int \int D^\gamma_x \left(\mathcal{Q}_{\emptyset,\xi} \left(e^{i(x,\xi)}\right)\right) \psi(x)\, dx f_F(\xi)\, d\xi + (2\pi)^{-\frac{d}{2}} \left[D^\gamma p_{\widehat{f}}, \psi\right] \\ &= (2\pi)^{-\frac{d}{2}} \int \int D^\gamma_x \left(\mathcal{Q}_{\emptyset,\xi} \left(e^{i(x,\xi)}\right)\right) f_F(\xi)\, \psi(x)\, dx d\xi + (2\pi)^{-\frac{d}{2}} \left[D^\gamma p_{\widehat{f}}, \psi\right].\end{aligned}$$

But from the assumptions of this theorem we know the double integral is absolutely integrable. Thus

$$[D^\gamma f, \psi] = (2\pi)^{-\frac{d}{2}} \int \left(\int D^\gamma_x \mathcal{Q}_{\emptyset,\xi} \left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi\right) \psi(x)\, dx + (2\pi)^{-\frac{d}{2}} \left[D^\gamma p_{\widehat{f}}, \psi\right],$$

and assumption 3.5 implies $\int D^\gamma_x \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi \in L^1_{loc}$, so we can conclude that

$$D^\gamma f(x) = (2\pi)^{-\frac{d}{2}} \int D^\gamma_x \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi + (2\pi)^{-\frac{d}{2}} D^\gamma p_{\widehat{f}}(x),$$

which proves 3.6. In addition,
$D^\gamma_x \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) f_F(\xi) \in C^{(0)}$ for each ξ, and by inequality 3.5,
$D^\gamma_x \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) f_F(\xi) \in L^1$ for each x. Thus by Lemma 107, $D^\gamma f \in C^{(0)}$,
and the upper bound 3.5 applied to the right side of 3.6 implies $D^\gamma f \in C^{(0)}_{BP}$.
Further, from 3.2

$$D^\gamma_x \mathcal{Q}_{\emptyset,n,\xi} \left(e^{i(x,\xi)}\right) = \begin{cases} (i\xi)^\gamma \mathcal{Q}_{\emptyset,n-|\gamma|,\xi} \left(e^{i(x,\xi)}\right), & |\gamma| < n, \\ (i\xi)^\gamma, & |\gamma| \geq n, \end{cases}$$

which proves 3.7.
In part 1 it was shown that $\widehat{D^\gamma f} = (i\xi)^\gamma f_F$ on $S_{\emptyset,n}$ and $(i\xi)^\gamma f_F \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$. Thus $\widehat{D^\gamma f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and it is meaningful to define the function $(D^\gamma f)_F$. In fact $(D^\gamma f)_F = (i\xi)^\gamma f_F$ a.e. on $\mathbb{R}^d \backslash 0$ and hence $(D^\gamma f)_F = (i\xi)^\gamma f_F$ a.e. which means the right side of 3.7 implies the right side of 3.8. ■

Remark 85 *Equation 3.3 can be regarded as a modified inverse Fourier transform for a subspace of the tempered distributions. Equation 3.6 can be regarded as a 'modified' inverse Fourier transform for a subspace of $C_{BP}^{(0)}$. Instead of using the Fourier transform as the argument its L_{loc}^1 restriction to $\mathbb{R}^d \backslash 0$ is used, and instead of the exponential $e^{i(x,\xi)}$, the exponential has a Taylor series-like term subtracted.*

For $\gamma = 0$ equation 3.6 can be written

$$\left(f - (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}\right)(x) = (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi,$$

and since $\left(f - (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}\right)_F = f_F$ the modified inverse Fourier transform can be thought of as only applying to functions of the form $f - (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}$.

The next two corollaries will be directly applicable to both basis functions and to functions belonging to X_w^θ.

Corollary 86 *Suppose m and n are positive integers. Further:*

1. *Suppose $f \in S'$ and $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$. Define a.e. the function $f_F : \mathbb{R}^d \to \mathbb{C}$ by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$.*
2. *Assume the action $\int f_F \phi$ on $\phi \in S_{\emptyset,n}$ defines a member of $S'_{\emptyset,n}$, and that $f_F = \widehat{f}$ on $S_{\emptyset,n}$.*
3. *Assume that if $|\gamma| \le m$ there exist $s_\gamma \ge 0$ and $k_\gamma > 0$ independent of ξ such that*

$$\int \left| D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) f_F(\xi)\right| d\xi \le k_\gamma (1+|x|)^{s_\gamma}. \tag{3.11}$$

Now define $f_\rho \in S'$ by

$$f = f_\rho + (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}, \tag{3.12}$$

where $p_{\widehat{f}} \in P_n$ is defined by 3.4 and the subscript $\rho \in S_{1,n}$ is the function used to define the operator $\mathcal{P}_{\emptyset,n}$.

Then $f_\rho \in C_{BP}^{(m)} \cap S'$ and if $|\gamma| \le m$ then

$$\begin{aligned} D^\gamma f_\rho(x) &= (2\pi)^{-\frac{d}{2}} \int D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi & (3.13)\\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right)(i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| < n,\\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| \ge n, \end{cases} & (3.14)\\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right)(D^\gamma f)_F(\xi)\, d\xi, & |\gamma| < n,\\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (D^\gamma f)_F(\xi)\, d\xi, & |\gamma| \ge n, \end{cases} & (3.15) \end{aligned}$$

and

$$|D^\gamma f_\rho(x)| \le k_\gamma (1+|x|)^{s_\gamma}, \quad |\gamma| \le m. \tag{3.16}$$

Proof. Equation 3.6 of Theorem 84 can be written

$$D^\gamma\left(f(x) - (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}(x)\right) = (2\pi)^{-\frac{d}{2}} \int D_x^\gamma \mathcal{Q}_{\emptyset,n,\xi}\left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi,$$

so the definition of f_ρ implies equations 3.13 to 3.8. Theorem 84 implies $D^\gamma f \in C^{(0)}$ for $|\gamma| < m$ i.e. $f \in C^{(m)}$. Further, the upper bound 3.11 applied to the right side of equation 3.13 implies the upper bound 3.16 and hence that $f_\rho \in C_{BP}^{(m)}$. ■

3.6 An inverse Fourier transform for basis functions

From Theorem 49 we see that when a weight function has property W3.1 the basis functions of order θ lie in $G \in C_B^{(\lfloor 2\kappa \rfloor)}$ i.e. the derivatives are bounded functions. However, for property W3.2 we only know that $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ i.e. G has polynomial growth at infinity but with no upper bound on the rate of growth near infinity. This situation will be rectified in Theorem 90 where we will use the upper bounds

for $\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|$ derived in the next lemma to give an upper bound for the basis function growth near infinity and to prove some inverse Fourier transform formulas for the basis functions and their derivatives. We then explore some other consequences of these results.

But first we need the lemma which will supply the estimates 3.5 required by Corollary 86.

Lemma 87 *Suppose the weight function w has properties W2.1 and W3.2 for order θ and parameter κ. Then if $|\gamma| \leq \lfloor 2\kappa \rfloor$ there exists a constant C_w, independent of x, such that*

$$\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|}{w(\xi)\,|\xi|^{2\theta}} d\xi \leq \begin{cases} C_w\,(1+|x|)^{2\theta-|\gamma|}, & |\gamma| < 2\theta, \\ C_w, & |\gamma| \geq 2\theta. \end{cases} \tag{3.17}$$

C_w is given by 3.20 and only depends on the weight function w, the function $\rho \in S_{1,2\theta}$ used to define $\mathcal{P}_{\emptyset,2\theta}$ and on the parameters θ, κ, r_3 which define weight function property W3.2.

Proof. There are two cases to be considered: $|\gamma| < 2\theta$ and $|\gamma| \geq 2\theta$. In both cases we will split the range of integration into the two concentric regions defined by the sphere $S(0;r_3)$.

We use the estimate of part 2 of Theorem 81 for $\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|$ inside the sphere $S(0;r_3)$ and the estimate of part 1 of Theorem 81 outside this sphere.

Case 1 $|\gamma| < 2\theta$ and $|\gamma| \leq \lfloor 2\kappa \rfloor$.

$$\begin{aligned}\int\limits_{|\xi|\leq r_3} \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|}{w(\xi)\,|\xi|^{2\theta}} d\xi &\leq \int\limits_{|\cdot|\leq r_3} \frac{C_{2\theta,\gamma,r_3}\,|\cdot|^{2\theta}\,(1+|x|)^{2\theta-|\gamma|}}{w\,|\cdot|^{2\theta}} \\ &\leq C_{2\theta,\gamma,r_3}\left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w}\right)(1+|x|)^{2\theta-|\gamma|} \\ &\leq \max_{|\lambda|<2\theta} C_{2\theta,\lambda,r_3}\left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w}\right)(1+|x|)^{2\theta-|\gamma|},\end{aligned}$$

and the integral exists since property W2.1 is $1/w \in L^1_{loc}$. Further

$$\begin{aligned}\int\limits_{|\xi|\geq r_3} \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|}{w(\xi)\,|\xi|^{2\theta}} d\xi &\leq \int\limits_{|\cdot|\geq r_3} \frac{C_{2\theta,\gamma}\,|\cdot|^{|\gamma|}\,(1+|x|)^{2\theta-|\gamma|-1}}{w\,|\cdot|^{2\theta}} \\ &\leq C_{2\theta,\gamma}\left(\int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}}\right)(1+|x|)^{2\theta-|\gamma|} \\ &\leq \max_{|\lambda|<2\theta} C_{2\theta,\lambda} \max_{|\lambda|\leq\lfloor 2\kappa\rfloor}\left(\int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\lambda|}}{w\,|\cdot|^{2\theta}}\right)(1+|x|)^{2\theta-|\gamma|},\end{aligned}$$

And by Theorem 11 the last integral exists. Adding the two estimates we get

$$\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)\right|}{w(\xi)\,|\xi|^{2\theta}} d\xi \leq C'_w\,(1+|x|)^{2\theta-|\gamma|}, \tag{3.18}$$

where

$$C'_w = \max\left\{\max_{|\lambda|<2\theta} C_{2\theta,\lambda},\ \max_{|\lambda|<2\theta} C_{2\theta,\lambda,r_3}\right\} \max_{|\lambda|\leq\lfloor 2\kappa\rfloor}\left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\lambda|}}{w\,|\cdot|^{2\theta}}\right). \tag{3.19}$$

Case 2 $|\gamma| \geq 2\theta$ and $|\gamma| \leq \lfloor 2\kappa \rfloor$. From Theorem 81

$$\begin{aligned}
\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\xi}\left(e^{i(x,\xi)}\right)\right|}{w(\xi)\,|\xi|^{2\theta}} &\leq \int\limits_{|\cdot|\leq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&= \int\limits_{|\cdot|\leq r_3} |\cdot|^{|\gamma|-2\theta}\,\frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&\leq r_3^{|\gamma|-2\theta} \int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&\leq (1+r_3)^{|\gamma|-2\theta} \left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\gamma|}}{w\,|\cdot|^{2\theta}} \right) \\
&\leq (1+r_3)^{2\kappa} \max_{|\lambda|\leq\lfloor 2\kappa\rfloor} \left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\lambda|}}{w\,|\cdot|^{2\theta}} \right).
\end{aligned}$$

We now combine both cases by setting

$$C_w = \max\left\{ \max_{|\lambda|<2\theta} C_{2\theta,\lambda},\ \max_{|\lambda|<2\theta} C_{2\theta,\lambda,r_3},\ (1+r_3)^{2\kappa} \right\} \max_{|\lambda|\leq\lfloor 2\kappa\rfloor} \left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\lambda|}}{w\,|\cdot|^{2\theta}} \right) \tag{3.20}$$

The constants $C_{2\theta,\gamma}$ and $C_{2\theta,\gamma,r_3}$ are defined in the proof of Theorem 81. ■

If we only assume the weight function has property W2 then part 1 of Theorem 84 allows us to convert the definition 1.33 of a basis distribution $G \in S'$ i.e. $\left[\widehat{G},\phi\right] = \int \frac{\phi}{w|\cdot|^{2\theta}}$ for all $\phi \in S_{\emptyset,2\theta}$, into the explicit formulas 3.21 of the next theorem.

Theorem 88 *Suppose the weight function w has property W2 and let G be a basis distribution of order $\theta \geq 1$ generated by w. Then for any multi-index γ*

$$\left[\widehat{D^\gamma G},\phi\right] = \int (i\xi)^\gamma \frac{\mathcal{Q}_{\emptyset,2\theta}\phi(\xi)}{w(\xi)\,|\xi|^{2\theta}}d\xi + (2\pi)^{-\frac{d}{2}} \left[\left(p_{\widehat{D^\gamma G}}\right)^\wedge,\phi\right], \quad \phi \in S, \tag{3.21}$$

where for $u \in S'$

$$p_u(x) = \sum_{|\alpha|<2\theta} \frac{b_{u,\alpha}}{\alpha!} x^\alpha, \quad b_{u,\alpha} = \left[u,(-i\xi)^\alpha \rho\right], \tag{3.22}$$

and ξ is the action variable.

Proof. From weight function property W2.1, $\frac{1}{w} \in L^1_{loc}$ and so by Definition 44 we have $G \in S'$, $\widehat{G} = \frac{1}{w|\cdot|^{2\theta}}$ on $S_{\emptyset,2\theta}$ and $\frac{1}{w|\cdot|^{2\theta}} \in L^1_{loc}\left(\mathbb{R}^d\backslash 0\right) \cap S'_{\emptyset,2\theta}$. Now set $G_F = \widehat{G}$ on $\mathbb{R}^d\backslash 0$ so that $G_F = \frac{1}{w|\cdot|^{2\theta}} \in L^1_{loc}\left(\mathbb{R}^d\backslash 0\right) \cap S'_{\emptyset,2\theta}$. Equation 3.21 now follows from equation 3.3 of Theorem 84 with $f = G$. ■

Remark 89 *This result is closely related to Theorem 2.1 of Madych and Nelson [13]. Indeed, in the comments following Theorem 2.1 Madych and Nelson illustrate Theorem 2.1 by choosing $d\mu(\xi) = w(\xi)\,d\xi$ where w corresponds to $1/w$ in this document i.e. to the **reciprocal** of our weight function.*

If we further assume that the weight function has property W3.2 then we can apply the bounds derived in the previous lemma to Corollary 86 and show that every basis function G is the sum of a polynomial in $P_{2\theta}$ and a special basis function G_ρ which depends only on the choice of a function $\rho \in S_{1,n}$. In turn G_ρ satisfies the modified inverse-Fourier transform equations 3.24 and the growth estimates 3.25.

Theorem 90 *Suppose the weight function w has property W2.1 and property **W3.2** for order θ and parameter κ. Further suppose that G is a basis distribution of order θ and the operator $\mathcal{P}_{\emptyset,\theta}$ is defined using the function $\rho \in S_{1,n}$.*

Then $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ *and we define the basis distribution* G_ρ *by*

$$G = G_\rho + (2\pi)^{-\frac{d}{2}} p_{\widehat{G}}, \tag{3.23}$$

where $p_{\widehat{G}} \in P_{2\theta}$ *is defined using 3.22. For all* $|\gamma| \leq \lfloor 2\kappa \rfloor$*:*

$$\begin{aligned} D^\gamma G_\rho(x) &= (2\pi)^{-\frac{d}{2}} \int \frac{D_x^\gamma \mathcal{Q}_{\emptyset,2\theta,\xi}\left(e^{i(x,\xi)}\right)}{w(\xi)|\xi|^{2\theta}} d\xi \\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,2\theta-|\gamma|,\xi}\left(e^{i(x,\xi)}\right) \frac{(i\xi)^\gamma}{w(\xi)|\xi|^{2\theta}} d\xi, & |\gamma| < 2\theta, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} \frac{(i\xi)^\gamma}{w(\xi)|\xi|^{2\theta}} d\xi, & |\gamma| \geq 2\theta, \end{cases} \\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,n-|\gamma|,\xi}\left(e^{i(x,\xi)}\right) (D^\gamma G_\rho)_F(\xi)\, d\xi, & |\gamma| < 2\theta, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (D^\gamma G_\rho)_F(\xi)\, d\xi, & |\gamma| \geq 2\theta, \end{cases} \end{aligned} \tag{3.24}$$

and $D^\gamma G_\rho$ *satisfies the growth estimate*

$$|D^\gamma G_\rho(x)| \leq \begin{cases} C_w (1+|x|)^{2\theta-|\gamma|}, & |\gamma| < 2\theta, \\ C_w, & |\gamma| \geq 2\theta. \end{cases} \tag{3.25}$$

The constant C_w *is given by 3.20.*

Proof. Since $p_{\widehat{G}} \in P_\theta$ the basis distribution Definition 44 implies G_ρ is also a basis distribution, and by part 4 of Theorem 50 it lies in $C_{BP}^{(\lfloor 2\kappa \rfloor)}$.

We now want to apply Corollary 86 with $f = G$. However, inspection of the proof of Theorem 88 and the inequality 3.25 shows that when $m = \lfloor 2\kappa \rfloor$, $n = 2\theta$, $f_F = \frac{1}{w|\cdot|^{2\theta}}$, $k_\gamma = C_w$ and $s_\gamma = 2\theta - |\gamma|$ for $|\gamma| \leq \lfloor 2\kappa \rfloor$, all the conditions of Corollary 86 are satisfied and we have our result. ■

The special basis function G_ρ defined by 3.23 of the last theorem will enable us to relate various weight function properties to the corresponding basis function properties e.g. the weight function is radial implies the function G_ρ is radial.

Lemma 91 *This lemma will employ the technique of Section 4.1, Stein and Weiss [5] which defines radial functions in terms of orthogonal transformations.*

Stein and Weiss observe that a function f *is radial if and only if* $f(\mathcal{O}x) = f(x)$ *for any linear, orthogonal transformation* $\mathcal{O} : \mathbb{R}^d \to \mathbb{R}^d$ *and any* $x \in \mathbb{R}^d$*.*

Note that an orthogonal transformation $\mathcal{O}$ *satisfies* $\mathcal{O}^T = \mathcal{O}^{-1}$ *where* $(\mathcal{O}x, y) = \left(x, \mathcal{O}^T y\right)$ *for the Euclidean inner product. An orthogonal transformation has a Jacobian of one.*

Theorem 92 *Suppose the weight function* w *has properties W2.1 and W3.2 for order* θ *and* κ*. Then the special basis function* G_ρ *of order* θ *introduced in Theorem 90 has the following properties:*

1. *If* ρ_1 *and* ρ_2 *are in* $S_{1,2\theta}$ *then* $\rho_2 - \rho_1 \in S_{\emptyset,2\theta}$ *and*

$$G_{\rho_1}(x) - G_{\rho_2}(x) = (2\pi)^{-d/2} \sum_{|\alpha|<2\theta} \left(\int \frac{\rho_2(\xi) - \rho_1(\xi)}{w(\xi)|\xi|^{2\theta}} (i\xi)^\alpha d\xi \right) \frac{x^\alpha}{\alpha!} \in P_{2\theta}.$$

2. $G_\rho(-x) = \overline{G_\rho(x)}$ *i.e.* G_ρ *is conjugate-even.*

3. *If* w *and* ρ *are even functions then* G_ρ *is a real valued even function.*

4. *If* w *and* ρ *are radial functions then* G_ρ *is also a radial function.*

Proof. Part 1 By part 5 of Theorem 81 the integrands defining G_{ρ_1} and G_{ρ_2} are absolutely integrable. Hence

$$\begin{aligned}(2\pi)^{d/2}\left(G_{\rho_1}(x)-G_{\rho_2}(x)\right)&=\int\left(e^{i(x,\xi)}-\rho_1(\xi)\sum_{|\alpha|<2\theta}\frac{(ix)^\alpha\xi^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}-\\&\quad-\int\left(e^{i(x,\xi)}-\rho_2(\xi)\sum_{|\alpha|<2\theta}\frac{x^\alpha(i\xi)^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}\\&=\int\left((\rho_2(\xi)-\rho_1(\xi))\sum_{|\alpha|<2\theta}\frac{x^\alpha(i\xi)^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}\\&=\int\sum_{|\alpha|<2\theta}\frac{\rho_2(\xi)-\rho_1(\xi)}{w(\xi)|\xi|^{2\theta}}\frac{x^\alpha(i\xi)^\alpha}{\alpha!}d\xi\\&=\sum_{|\alpha|<2\theta}\left(\int\frac{\rho_2(\xi)-\rho_1(\xi)}{w(\xi)|\xi|^{2\theta}}(i\xi)^\alpha d\xi\right)\frac{x^\alpha}{\alpha!},\end{aligned}$$

if the integrals on the last line all exist. However, by parts 3 and 4 of Theorem 15, $\rho_2(\xi)-\rho_1(\xi)\in S_{\emptyset,2\theta}$ implies $(\rho_2(\xi)-\rho_1(\xi))(i\xi)^\alpha\in S_{\emptyset,2\theta}$. But Theorem 42 showed that if $\phi\in S_{\emptyset,2\theta}$ then $\int\int\frac{|\phi(\xi)|}{w(\xi)|\xi|^{2\theta}}<\infty$ and so $\frac{\rho_2(\xi)-\rho_1(\xi)}{w(\xi)|\xi|^{2\theta}}(i\xi)^\alpha\in L^1$.

Part 2

$$\overline{G_\rho(x)}=(2\pi)^{-\frac{d}{2}}\int\left(e^{-i(x,\xi)}-\rho(\xi)\sum_{|\alpha|<2\theta}\frac{(-ix)^\alpha\xi^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}=G_\rho(-x).$$

Part 3 Suppose w and ρ are even functions. Then $\rho(-x)=\rho(x)$ and

$$\begin{aligned}(2\pi)^{\frac{d}{2}}G_\rho(-x)&=\int\left(e^{-i(x,\xi)}-\rho(\xi)\sum_{|\alpha|<2\theta}\frac{(-ix)^\alpha\xi^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}\\&=\int\left(e^{-i(x,\xi)}-\rho(-\xi)\sum_{|\alpha|<2\theta}\frac{(ix)^\alpha(-\xi)^\alpha}{\alpha!}\right)\frac{d\xi}{w(-\xi)|\xi|^{2\theta}}\\&=\int\left(e^{i(x,\xi)}-\rho(\xi)\sum_{|\alpha|<2\theta}\frac{(ix)^\alpha\xi^\alpha}{\alpha!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}\\&=(2\pi)^{d/2}G_\rho(x).\end{aligned}$$

Part 2 implies G_ρ is real.

Part 4 First note that $S_{1,2\theta}$ contains radial functions e.g. the standard example of a distribution test function, the 'cap-shaped' function

$$\begin{cases}e\exp\left(-\frac{1}{1-|x|^2}\right), & |x|\le 1,\\0, & |x|>1,\end{cases}$$

lies in $S_{1,\infty}=\{\phi\in S:\phi(0)=1,\ D^\alpha\phi(0)=0\ for\ all\ \alpha\neq 0\}$.

We now use the definition of a radial function given in Lemma 91 to prove that G_{ρ_r} is radial if ρ_r is radial. Now

$$G_{\rho_r}(x)=(2\pi)^{-\frac{d}{2}}\int\left(e^{i(x,\xi)}-\rho_r(\xi)\sum_{k=0}^{2\theta}\frac{i^k(x,\xi)^k}{k!}\right)\frac{d\xi}{w(\xi)|\xi|^{2\theta}}.$$

Thus

$$G_{\rho_r}(\mathcal{O}x) = (2\pi)^{-\frac{d}{2}} \int \left(e^{i(\mathcal{O}x,\xi)} - \rho_r(\xi) \sum_{k=0}^{2\theta} \frac{i^k (\mathcal{O}x,\xi)^k}{k!} \right) \frac{d\xi}{w(\xi)\,|\xi|^{2\theta}}$$
$$= (2\pi)^{-\frac{d}{2}} \int \left(e^{i\left\langle x,\mathcal{O}^T\xi \right\rangle} - \rho_r(\xi) \sum_{k=0}^{2\theta} \frac{i^k \left(x,\mathcal{O}^T\xi\right)^k}{k!} \right) \frac{d\xi}{w(\xi)\,|\xi|^{2\theta}},$$

and then making the change of variables $\eta = \mathcal{O}^T\xi$ with $d\eta = d\xi$ and $\xi = \mathcal{O}\eta$

$$G_{\rho_r}(\mathcal{O}x) = (2\pi)^{-\frac{d}{2}} \int \left(e^{i\langle x,\eta\rangle} - \rho_r(\mathcal{O}\eta) \sum_{k=0}^{2\theta} \frac{i^k (x,\eta)^k}{k!} \right) \frac{d\eta}{w(\mathcal{O}\eta)\,|\mathcal{O}\eta|^{2\theta}}$$
$$= (2\pi)^{-\frac{d}{2}} \int \left(e^{i\langle x,\eta\rangle} - \rho_r(\eta) \sum_{k=0}^{2\theta} \frac{i^k (x,\eta)^k}{k!} \right) \frac{d\eta}{w(\mathcal{O}\eta)\,|\mathcal{O}\eta|^{2\theta}}$$
$$= G_{\rho_r}(x),$$

since ρ_r and w are radial. Hence G_{ρ_r} is also radial. ■

Corollary 93 *Assume the weight function w has properties W2.1 and W3.2 for some order θ and parameter κ. Then:*

1. *There always exists a conjugate-even basis function.*
2. *If w is even then there exists an even, real valued basis function.*
3. *If w is radial then there exists a radial basis function.*

3.7 An inverse Fourier transform for data functions

We now recall some properties of the semi-Hilbert data spaces X_w^θ (Definition 21) needed in this section. These results are taken from Section 1.4 and the relevant theorem is given in brackets.

Summary 94 *Suppose w is a weight function with property W2. If $f \in X_w^\theta$ then $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and we can define a.e. the function $f_F : \mathbb{R}^d \to \mathbb{C}$ by: $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Further:*

1. *The seminorm and semi-inner product are given by*

$$\int w\,|\cdot|^{2\theta} f_F \overline{g_F} = \langle f,g \rangle_{w,\theta}, \quad \int w\,|\cdot|^{2\theta} |f_F|^2 = |f|^2_{w,\theta}. \tag{3.26}$$

   An alternative definition of X_w^θ is

$$X_w^\theta = \left\{ f \in S' : \xi^\alpha \widehat{f} \in L^1_{loc} \ \mathit{if}\ |\alpha| = \theta;\ \int w\,|\cdot|^{2\theta} |f_F|^2 < \infty \right\}. \tag{3.27}$$

   Note that by part 2 Lemma 24 the condition $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ is actually implied by $\xi^\alpha \widehat{f} \in L^1_{loc}$ for all $|\alpha| = \theta$.

2. *The functional $|\cdot|_{w,\theta}$ is a seminorm. In fact, $\mathrm{null}\,|\cdot|_{w,\theta} = P_\theta$ and $P \cap X_w^\theta = P_\theta$ (part 3 Theorem 25).*

3. *$f_F \in S'_{\emptyset,\theta}$ with action $\int f_F \phi$ on $S_{\emptyset,\theta}$. Also $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$ (part 2 Theorem 27).*

4. *X_w^θ is complete in the seminorm sense (Theorem 37).*

5. *If w has properties W2.1 and W3 for order θ and smoothness parameter κ then $X_w^\theta \subset C_{BP}^{(\lfloor\kappa\rfloor)}$ (Theorem 40).*

If a weight function w has properties W2 and W3 for order θ and smoothness parameter κ, we know from part 6 of Summary 94 that $X_w^\theta \subset C_{BP}^{(\lfloor\kappa\rfloor)}$ i.e. functions in X_w^θ have polynomial growth. The next lemma will allow us to use Corollary 86 to prove an inverse Fourier transform result for the functions in X_w^θ and to derive upper bounds for the growth rate of the derivatives near infinity.

Lemma 95 *Suppose the weight function w has properties W2.1 and W3 for order θ and smoothness parameter κ. Then if $|\gamma| \leq \lfloor\kappa\rfloor$ there exists a constant C_w, independent of x, such that*

$$\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2}{w(\xi)\,|\xi|^{2\theta}} d\xi \leq \begin{cases} (C_w)^2\,(1+|x|)^{2(\theta-|\gamma|)}, & |\gamma| < \theta, \\ (C_w)^2, & |\gamma| \geq \theta. \end{cases} \tag{3.28}$$

C_w is given by 3.32 and only depends on the weight function w, the function $\rho \in S_{1,2\theta}$ used to define $\mathcal{Q}_{\emptyset,\theta}$ and on the parameters which define the weight function properties W3.

Proof. There are two cases to be considered: $|\gamma| < \theta$ and $|\gamma| \geq \theta$. Define the constant r_4 by $r_4 = 0$ if w has property W3.1 and by $r_4 = r_3$ if w has property W3.2. In both cases we will split the range of integration into the two concentric regions defined by the sphere $S(0;r_4)$. We use the estimate of part 2 of Theorem 81 for $\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|$ inside the sphere $S(0;r_4)$ and the estimate of part 1 of Theorem 81 outside this sphere.

Case 1 $|\gamma| < \theta$ and $|\gamma| \leq \lfloor\kappa\rfloor$.

$$\begin{aligned}
\int\limits_{|\xi|\leq r_4} \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2 d\xi}{w(\xi)\,|\xi|^{2\theta}} &\leq \int\limits_{|\cdot|\leq r_4} \frac{(C_{\theta,\gamma,r_4})^2\,|\cdot|^{2\theta}\,(1+|x|)^{2(\theta-|\gamma|)}}{w\,|\cdot|^{2\theta}} \\
&= (C_{\theta,\gamma,r_4})^2 \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w}\right) (1+|x|)^{2(\theta-|\gamma|)} \\
&\leq \left(\max_{|\lambda|<\theta} C_{\theta,\lambda,r_4}\right)^2 \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w}\right) (1+|x|)^{2(\theta-|\gamma|)},
\end{aligned}$$

which exists since property W2.1 is $1/w \in L^1_{loc}$. Further

$$\begin{aligned}
\int\limits_{|\xi|\geq r_4} \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2 d\xi}{w(\xi)\,|\xi|^{2\theta}} &\leq \int\limits_{|\cdot|\geq r_4} \frac{(C_{\theta,\lambda})^2\,|\cdot|^{2|\gamma|}\,(1+|x|)^{2(\theta-|\gamma|-1)}}{w\,|\cdot|^{2\theta}} \\
&= (C_{\theta,\lambda})^2 \left(\int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}}\right) (1+|x|)^{2(\theta-|\gamma|)} \\
&\leq \left(\max_{|\lambda|<\theta} C_{\theta,\lambda}\right)^2 \left(\max_{|\lambda|<\lfloor\kappa\rfloor} \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\lambda|}}{w\,|\cdot|^{2\theta}}\right) (1+|x|)^{2(\theta-|\gamma|)}
\end{aligned}$$

and since w has property W3, the definition of r_4 and Theorem 11 implies the last integral exists. Adding the last two estimates we get

$$\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2}{w(\xi)\,|\xi|^{2\theta}} d\xi \leq (C'_w)^2\,(1+|x|)^{2(\theta-|\gamma|)}, \tag{3.29}$$

where

$$C'_w = \max\left\{\max_{|\lambda|<\theta} C_{\theta,\lambda}, \max_{|\lambda|<\theta} C_{\theta,\lambda,r_4}\right\} \max_{|\lambda|\leq\lfloor\kappa\rfloor} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\lambda|}}{w\,|\cdot|^{2\theta}}\right)^{\frac{1}{2}}. \tag{3.30}$$

Case 2 $|\gamma| \geq \theta$ and $|\gamma| \leq \lfloor\kappa\rfloor$

$$\begin{aligned}
\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2}{w(\xi)\,|\xi|^{2\theta}} d\xi &= \int\limits_{|\cdot|\leq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&= \int\limits_{|\cdot|\leq r_4} |\cdot|^{2(|\gamma|-\theta)} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&\leq r_4^{2(|\gamma|-\theta)} \int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}} \\
&\leq (1+r_4)^{2(|\gamma|-\theta)} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}} \right) \\
&\leq (1+r_4)^{2\kappa} \max_{|\lambda|\leq\lfloor\kappa\rfloor} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{|\cdot|^{2|\lambda|}}{w\,|\cdot|^{2\theta}} \right). \qquad (3.31)
\end{aligned}$$

Since property W2.1 requires that $1/w \in L^1_{loc}$and property W3 implies Theorem 11, it follows that the last two integrals exist. We now combine both cases by setting

$$C_w = \max\left\{ \max_{|\lambda|<\theta} C_{\theta,\lambda}, \max_{|\lambda|<\theta} C_{\theta,\lambda,r_4}, (1+r_4)^{2\kappa} \right\} \max_{|\lambda|\leq\lfloor\kappa\rfloor} \left(\int\limits_{|\cdot|\leq r_3} \frac{1}{w} + \int\limits_{|\cdot|\geq r_3} \frac{|\cdot|^{|\lambda|}}{w\,|\cdot|^{2\theta}} \right) \qquad (3.32)$$

so that inequalities 3.29 and 3.31 imply the

The constants $C_{\theta,\lambda}$ and C_{θ,λ,r_4} are defined in the proof of Theorem 81. ∎

The next theorem is our inverse Fourier transform result for distributions in X_w^θ when the weight function has property W2.

Theorem 96 *Suppose the weight function w has property W2 and suppose $f \in X_w^\theta$. Summary 94 allows us to define a.e. the function $f_F : \mathbb{R}^d \to \mathbb{C}$ by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Then for all multi-indexes γ*

$$\left[\widehat{D^\gamma f}, \psi\right] = \int (i\xi)^\gamma \mathcal{Q}_{\emptyset,\theta}\psi(\xi)\, f_F(\xi)\, d\xi + (2\pi)^{-\frac{d}{2}} \left[\widehat{D^\gamma p_{\widehat{D^\gamma f}}}, \psi\right], \quad \psi \in S, \qquad (3.33)$$

where for $u \in S'$, $p_u \in P_\theta$ and

$$p_u(x) = \sum_{|\alpha|<\theta} \frac{b_{u,\alpha}}{\alpha!} x^\alpha, \quad b_{u,\alpha} = \left[u, (-i\xi)^\alpha \rho\right]. \qquad (3.34)$$

Proof. If $f \in X_w^\theta$ then by Summary 94 $f \in S'$, $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$, $f_F \in S'_{\emptyset,\theta}$ with action$\int f_F \phi$ and $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$. Hence f satisfies the assumptions of part 1 of Theorem 84 and 3.3 and 3.4 implies 3.33 and 3.34. ∎

If a weight function also has property W3 for order θ then we have the following modified inverse-Fourier transform theorem for the (continuous) functions in X_w^θ as well as an upper bound for the growth rate 3.37 of the derivatives near infinity. More precisely, a function in X_w^θ is shown to be the sum of a function $f_\rho \in X_w^\theta$ and a polynomial in P_θ and the function f_ρ satisfies the modified inverse-Fourier transform equations 3.36.

Theorem 97 *Now suppose the weight function w also has properties W2 and W3 for order θ and smoothing parameter κ. Then $X_w^\theta \subset C_{BP}^{(\lfloor\kappa\rfloor)}$. Now define the function $f_\rho \in X_w^\theta$ by*

$$f = f_\rho + (2\pi)^{-\frac{d}{2}} p_{\widehat{f}}, \quad f \in X_w^\theta, \qquad (3.35)$$

where $p_{\widehat{f}} \in P_\theta$ is defined by 3.34 and $\rho \in S_{1,n}$ is used to define $\mathcal{P}_{\emptyset,\theta}$. Then $f_\rho = \mathcal{Q}_{\emptyset,\theta} f \in C_{BP}^{(\lfloor\kappa\rfloor)} \cap X_w^\theta$ and for $|\gamma| \leq \kappa$:

$$\begin{aligned} D^\gamma f_\rho(x) &= (2\pi)^{-\frac{d}{2}} \int D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right) f_F(\xi)\, d\xi \\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,\theta-|\gamma|,\xi}\left(e^{i(x,\xi)}\right)(i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| < \theta, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (i\xi)^\gamma f_F(\xi)\, d\xi, & |\gamma| \geq \theta, \end{cases} \\ &= \begin{cases} (2\pi)^{-\frac{d}{2}} \int \mathcal{Q}_{\emptyset,\theta-|\gamma|,\xi}\left(e^{i(x,\xi)}\right)(D^\gamma f_\rho)_F(\xi)\, d\xi, & |\gamma| < \theta, \\ (2\pi)^{-\frac{d}{2}} \int e^{i(x,\xi)} (D^\gamma f_\rho)_F(\xi)\, d\xi, & |\gamma| \geq \theta, \end{cases} \end{aligned} \tag{3.36}$$

and $D^\gamma f_\rho$ satisfies the growth estimates

$$|D^\gamma f_\rho(x)| \leq \begin{cases} C_w |f_\rho|_{w,\theta} (1+|x|)^{\theta-|\gamma|}, & |\gamma| < \theta, \\ C_w |f_\rho|_{w,\theta}, & |\gamma| \geq \theta, \end{cases} \tag{3.37}$$

where the constant C_w was defined in Lemma 95.

Proof. This proof is an application of Corollary 86 with $n = \theta$ and $m = \lfloor\kappa\rfloor$. If $f \in X_w^\theta$ then by Summary 94, $f \in S'$, $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$, $f_F \in S'_{\emptyset,\theta}$ with action$\int f_F \phi$ and $f_F = \widehat{f}$ on $S_{\emptyset,\theta}$ so f satisfies assumptions 1 and 2 of Corollary 86. Regarding assumption 3: using the Cauchy-Schwartz inequality

$$\begin{aligned} \int \left|\left(D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right) f_F(\xi)\right| d\xi &= \int \left| \frac{D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)}{\sqrt{w(\xi)}\,|\xi|^\theta} \sqrt{w(\xi)}\,|\xi|^\theta f_F(\xi) \right| d\xi \\ &\leq \left(\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2}{w(\xi)\,|\xi|^{2\theta}} d\xi \right)^{\frac{1}{2}} |f|_{w,\theta}, \end{aligned}$$

since $|f|_{w,\theta}^2 = \int w\, |\cdot|^{2\theta} |f_F|^2\, d\xi$ by part 1 Summary 94. Applying the inequality derived in Lemma 95 for $|\gamma| \leq \lfloor\kappa\rfloor$ we get

$$\left(\int \frac{\left|D_x^\gamma \mathcal{Q}_{\emptyset,\theta,\xi}\left(e^{i(x,\xi)}\right)\right|^2}{w(\xi)\,|\xi|^{2\theta}} d\xi \right)^{\frac{1}{2}} \leq \begin{cases} C_w |f|_{w,\theta} (1+|x|)^{\theta-|\gamma|}, & |\gamma| < \theta, \\ C_w |f|_{w,\theta}, & |\gamma| \geq \theta. \end{cases}$$

Inequality 3.16 of Corollary 86 is now satisfied for $k_\gamma = C_w |f|_{w,\theta}$ and $s_\gamma = \max\{0, \theta - |\gamma|\}$ and so the rest of Corollary 86 can be applied to obtain $f \in C_{BP}^{(\lfloor\kappa\rfloor)}$ and all equations and estimates for $D^\gamma f_\rho$ except equation 3.36 where f_ρ has replaced f. But by 3.35, $D^\gamma f = D^\gamma f_\rho + (2\pi)^{-\frac{d}{2}} D^\gamma p_{\widehat{f}}$ and since $D^\gamma p_{\widehat{f}}$ is a polynomial we have $(D^\gamma f_\rho)_F = (D^\gamma f)_F$. ■

Remark 98 *This result is closely related to Proposition 2.2 of Madych and Nelson [13]. However, in the comments following Theorem 2.1 Madych and Nelson illustrate Theorem 2.1 by choosing $d\mu(\xi) = w(\xi)\, d\xi$ where w corresponds to $1/w$ in this document i.e. to the **reciprocal** of our weight function.*

4

The basis function interpolant and its convergence to the data function

4.1 Introduction

The goal of this chapter is to derive orders for the pointwise convergence of an interpolant to its data function as the density of the independent data increases.

Section by section:

Section 4.2 The concept of unisolvent data sets and minimal unisolvent data sets is introduced together with the Lagrangian interpolation operator $\mathcal{P}$ and the operator $\mathcal{Q} = I - \mathcal{P}$ which are defined using a minimal unisolvent subset of the data.

Section 4.3 Smoothness results and upper bounds are obtained for the operators $D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\xi)} \right)$. These are then used to estimate an upper bound for $\int \frac{\left| D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\xi)} \right) \right|^2}{w(\xi) |\xi|^{2\theta}} d\xi$ where w is the weight function.

Section 4.4 The properties of the function $\mathcal{Q}_x \left(e^{i(x,\xi)} \right)$ are used to derive an 'inverse Fourier transform' theorem which expresses the value of a function $f \in X_w^\theta$ in terms of its distribution Fourier transform, which is a function in $L_{loc}^1 \left(\mathbb{R}^d \backslash 0 \right)$.

Section 4.5 The data function space X_w^θ is endowed with the *Light norm* which is based on a minimal unisolvent set. We then derive explicit formulas for the Riesz representers of the evaluation functionals $f \to D^\gamma f(x)$ where $f \in X_w^\theta$. These formulas are expressed in terms of a basis function. The existence of a Riesz representer of $f \to f(x)$ means that X_w^θ is a reproducing kernel Hilbert space of continuous functions.

Section 4.6 In this section we define the *unisolvency matrix* and the *cardinal unisolvency matrix*, as well as the *basis function matrix* and the *reproducing kernel matrix*. These matrices will be used to construct the matrix equations for the interpolant as well as the smoothers studied in later chapters.

Section 4.7 The finite dimensional space $W_{G,X}$ is introduced and some properties proved. This space will contain the solution to the interpolation problem of this chapter and to the smoothing problems of later chapters.

Section 4.8 The vector-valued evaluation operator $\widetilde{\mathcal{E}}_X f = \left(f \left(x^{(k)} \right) \right)$ is studied. This operator and its adjoint will be fundamental to solving the interpolation and the smoothing problems.

Section 4.9 The minimal seminorm and norm interpolants are defined and shown to have the same basis function solution in $W_{G,X}$. The concept of a data function is introduced and several matrix equations for the smoother are derived.

Section 4.10 We obtain estimates for the order of pointwise convergence of the interpolant to its data function.

Section 4.11 The convergence estimates of the previous section are improved using Taylor series expansions of distributions. These results are applied to the thin-plate splines and the shifted thin-plate splines.

I have not included the results of any numerical experiments concerning the interpolant error estimates.

4.2 Unisolvency: sets, bases, operators and matrices

The concept of unisolvent sets of data is fundamental to the theory of basis function interpolation for orders $\theta \geq 1$. The basic importance of unisolvency is that it ensures that the interpolation problem has a unique solution. Using minimal unisolvent sets we will then construct the Lagrange polynomial interpolation operator $\mathcal{P}$ and the operator $\mathcal{Q} = I - \mathcal{P}$ as well as the unisolvency matrices.

In Subsection 4.5.2, following Light and Wayne [11], a unisolvent set of points will be used to define an inner product on the space X_w^θ. With this inner product we will show that X_w^θ is a reproducing kernel Hilbert space. But first we need to define unisolvency and some related concepts.

Definition 99 ***Unisolvent sets and minimal unisolvent sets***
Recall that P_θ is the set of polynomials of order θ i.e. of degree $\theta - 1$ when $\theta \geq 1$.

Then a finite (ordered) set of distinct points $X = \{x_i\}$ is said to be a ***unisolvent set*** *with respect to P_θ if: $p \in P_\theta$ and $p(x_i) = 0$ for all $x_i \in X$ implies $p = 0$.*

Sometimes we say X is unisolvent of order θ or that X is θ-unisolvent.

It is well known that any unisolvent set has at least $M = \dim P_\theta$ points, and that any unisolvent set of more than M points has a unisolvent set with M points. Consequently, a unisolvent set with M points is called a ***minimal unisolvent set****.*

It is convenient to introduce cardinal bases for polynomials and permutations together.

Definition 100 ***Cardinal bases for polynomials*** *and* ***permutations****.*

1. *A basis $\{l_i\}_{i=1}^M$ for P_θ is a* ***cardinal basis*** *for the minimal unisolvent set $A = \{a_i\}_{i=1}^M$ if $l_i(a_j) = \delta_{i,j}$ and the l_i are polynomials with* ***real valued*** *coefficients.*

2. *We will call the column vector $\widetilde{l} = (l_i)$ the* ***cardinal basis (column) vector*** *and we will write $\widetilde{l}_A$ to make the dependency on A explicit. Evaluation at a point x will be indicated by $\widetilde{l}_A(x)$.*

3. *It is well known that a set is minimally unisolvent iff there exists a (unique) cardinal basis for the set. Also, re-ordering A re-orders $\{l_i\}_{i=1}^M$ identically i.e. $\widetilde{l}_{\pi(A)} = \pi\left(\widetilde{l}_A\right) = \Pi\widetilde{l}_A$ where π is a permutation and Π is the corresponding permutation matrix. A permutation π can also be thought of as a sequence of indexes such that if $\left\{y^{(i)}\right\} = \pi(X)$ then $y^{(i)} = x^{(\pi(i))}$.*

4. *When used as a vector an ordered set is regarded as a column vector by default. Re-ordering by columns involves left-multiplication by Π. Re-ordering a row vector involves right-multiplying by Π^T.*

5. *The inverse permutation is denoted π^{-1} and has permutation matrix Π^T. Thus $\Pi^T\Pi = \Pi\Pi^T = I$.*

4.2.1 The Lagrange interpolation operator $\mathcal{P}$ and $\mathcal{Q} = I - \mathcal{P}$

Now we have conditions for the continuity of functions in X_w^θ, we can introduce operators, norms etc. which involve evaluating these functions at points. Following Light and Wayne, the first step is to introduce the operators $\mathcal{P}$ and $\mathcal{Q}$, which will play a pivotal role in the remainder of this document. Light and Wayne [11] introduced the Lagrange polynomial interpolation operator $\mathcal{P}$ as equation (7) following Lemma 3.4 as well as the operator $\mathcal{Q} = I - \mathcal{P}$.

Definition 101 ***The operators $\mathcal{P} : C^{(0)} \to P_\theta$ and $\mathcal{Q} : C^{(0)} \to C^{(0)}$. $\mathcal{P}$ is the Lagrange interpolation operator****.*
Suppose the set $A = \{a_i\}_{i=1}^M$ is a minimal unisolvent set with respect to the polynomials P_θ. Suppose $\{l_i\}_{i=1}^M$ is the cardinal basis of P_θ with respect to the unisolvent set of points A.

Then for any continuous function f define the operators $\mathcal{P}$ and $\mathcal{Q}$ by

$$\mathcal{P}f = \sum_{i=1}^{M} f(a_i)\, l_i, \qquad \mathcal{Q}f = f - \mathcal{P}f, \tag{4.1}$$

where the unisolvent set can be used as a subscript when necessary.

Theorem 102 *The operators $\mathcal{P}$ and $\mathcal{Q}$ have the following properties:*

1. *$p \in P_\theta$ implies $\mathcal{P}p = p$ and hence $\mathcal{Q}p = 0$.*

2. *For all $a_i \in A$, $(\mathcal{P}f)(a_i) = f(a_i)$ and $(\mathcal{Q}f)(a_i) = 0$. The polynomial $\mathcal{P}f$ interpolates the data $\{\ (a_i, f(a_i))\}_{i=1}^{M}$.*

3. *$\mathcal{P}^2 = \mathcal{P}$, $\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P} = 0$ and $\mathcal{Q}^2 = \mathcal{Q}$, so $\mathcal{P}$ and $\mathcal{Q}$ are projections.*

4. *$\operatorname{null}\mathcal{Q} = \operatorname{range}\mathcal{P} = P_\theta$.*

5. *$\mathcal{P}$ and $\mathcal{Q}$ are independent of the order of A.*

Proof. Part 1 is true since each member of the cardinal basis satisfies $\mathcal{P}l_j = \sum\limits_{i=1}^{M} l_i(a_i)\, l_i = l_j$ and **part 2** is true since $l_i(a_j) = \delta_{i,j}$. Regarding **part 3**, $\mathcal{P}$ is a projection since

$$\mathcal{P}^2 f = \mathcal{P}\left(\sum_{j=1}^{M} f(a_j)\, l_j\right) = \sum_{j=1}^{M} f(a_j)\, \mathcal{P}(l_j) = \sum_{j=1}^{M} f(a_j)\, l_j = \mathcal{P}f.$$

Thus $\mathcal{P}\mathcal{Q} = \mathcal{P}(I - \mathcal{P}) = \mathcal{P} - \mathcal{P}^2 = 0 = (I - \mathcal{P})\mathcal{P} = \mathcal{Q}\mathcal{P}$, and so $\mathcal{Q}^2 = \mathcal{Q}(I - \mathcal{P}) = \mathcal{Q}$. Finally **part 4** is true since $\mathcal{P} + \mathcal{Q} = I$. **Part 5** is true since from the definition of unisolvency reordering A reorders $\{l_i\}_{i=1}^{M}$ identically. In terms of permutations (recall Definition 100), if we write $\mathcal{P}_A$ and define the *vector-valued evaluation function* by $\widetilde{\mathcal{E}}_A f = (f(a_i))$ then

$$\begin{aligned}\mathcal{P}_{\pi(A)} f(x) = \left(\widetilde{\mathcal{E}}_{\pi(A)} f\right)^T \widetilde{l}_{\pi(A)}(x) = \left(\Pi \widetilde{\mathcal{E}}_A f\right)^T \Pi \widetilde{l}_A(x) = \left(\widetilde{\mathcal{E}}_A f\right)^T \Pi^T \Pi \widetilde{l}_A(x) &= \left(\widetilde{\mathcal{E}}_A f\right)^T \widetilde{l}_A(x) \\ &= \mathcal{P}_A f(x).\end{aligned}$$

■

4.2.2 Unisolvency matrices

This section builds on Section 4.2 which introduced the concepts of unisolvent sets, minimal unisolvent sets and cardinal bases. In this section we define the unisolvency matrix and the cardinal unisolvency matrix which can be used to characterize unisolvent sets in a form suitable for studying the interpolation and smoothing problems of this series of documents.

Definition 103 ***Unisolvency matrix*** *and* ***cardinal unisolvency matrix****.*

Suppose $\{p_i\}_{i=1}^{M}$ is any basis for the polynomials P_θ and we have an arbitrary ordered set of distinct points $X = \left\{x^{(i)}\right\}_{i=1}^{N} \subset \mathbb{R}^d$. Clearly we must have $N \geq M = \dim P_\theta$.

*The $N \times M$ **unisolvency matrix** P_X is now defined by*

$$P_X = \left(p_j\left(x^{(i)}\right)\right). \tag{4.2}$$

For example, for numerical work we may use the monomials $\{x^\alpha\}_{|\alpha|<\theta}$ as a basis for P_θ.

Now suppose the subset $A \subset X$ is minimally unisolvent and has cardinal basis $\{l_j\}_{j=1}^{M}$. Then we use the special notation

$$L_X = \left(l_j\left(x^{(i)}\right)\right), \tag{4.3}$$

*for the unisolvency matrix which corresponds to the basis $\{l_j\}$. We call L_X a **cardinal unisolvency matrix**.*

As well as proving some elementary properties of unisolvency matrices, the next theorem introduces some matrix notation which will be used often later. It will be noted that there is an emphasis on characterizing the set $\operatorname{null} P_X^T$. This is because the constraint $P_X^T v = 0$ is part of the definition (Definition 130) of the finite dimensional space $W_{G,X}$ which contains the interpolants studied in this document. We now prove two theorems of properties of the unisolvency matrices. The first theorem proves results which are true for any minimal unisolvent subset of X. The second theorem assumes the first M points of X constitute a minimal unisolvent subset.

Theorem 104 ***Properties of unisolvency matrices*** *Suppose $X = \{x^{(i)}\}_{i=1}^{N}$ is any set of distinct points in $\mathbb{R}^d$. Then if $M = \dim P_\theta$:*

1. *The null space of the unisolvency matrix P_X can be used to characterize unisolvent sets. In fact, X is unisolvent if and only if $\operatorname{null} P_X = \{0\}$.*

   Now suppose the set X is unisolvent and $A = \{a_i\}_{i=1}^{M}$ is any minimal unisolvent subset. Then:

2. *$P_X^T \beta = 0$ iff $\sum\limits_{j=1}^{N} \beta_j p\left(x^{(j)}\right) = 0$ for all $p \in P_\theta$.*

3. *Suppose P_X and Q_X are the unisolvency matrices generated by the bases $\widetilde{p} = \{p_i\}_{i=1}^{M}$ and $\widetilde{q} = \{q_i\}_{i=1}^{M}$ of P_θ. Let R be the regular change of basis matrix i.e. $\widetilde{p} = R\widetilde{q}$. Then*
$$P_X = Q_X R^T = Q_X Q_A^{-1} P_A,$$
   where $P_A = (p_j(a_i))_{i,j=1}^{M}$ and $Q_A = (q_j(a_i))_{i,j=1}^{M}$ are regular. Further
$$P_X = L_X P_A. \tag{4.4}$$

4. *$P_X^T \beta = 0$ iff $L_X^T \beta = 0$.*

5. *$\dim \operatorname{null} P_X^T = \dim \operatorname{null} L_X^T = N - M$.*

Proof. Part 1 If $p \in P_\theta$ then $p = \sum\limits_{i=1}^{M} \beta_i p_i$ for unique β_i. Thus the condition $p\left(x^{(j)}\right) = 0$ for all $j = 1, \ldots, N$ is equivalent to $\sum\limits_{i=1}^{M} \beta_i p_i\left(x^{(j)}\right) = 0$ for all $j = 1, \ldots, N$ which by definition 4.2 is equivalent to $P_X \beta = 0$.

Part 2 Since $0 = P_X^T \beta = \sum\limits_{j=1}^{N} \beta_j p_i\left(x^{(j)}\right)$ and $\{p_i\}_{i=1}^{M}$ is a basis for P_θ, it follows that $\sum\limits_{j=1}^{N} \beta_j p\left(x^{(j)}\right) = 0$ for any $p \in P_\theta$. Clearly the argument is reversible.

Part 3 If $\widetilde{p} = (p_i)$ and $\widetilde{q} = (q_i)$ are the basis vectors then there is a regular change of basis matrix R such that $\vec{p} = R\,\widetilde{q}$. Then $(\widetilde{p})^T = (\widetilde{q})^T R^T$ so that $(\widetilde{p})^T \left(x^{(j)}\right) = (\widetilde{q})^T \left(x^{(j)}\right) R^T$ i.e. $P_X = Q_X R^T$ which implies $P_A = Q_A R^T$ and so $P_X = Q_X R^T = Q_X Q_A^{-1} P_A$. If we chose $(\widetilde{q})$ to be the cardinal basis i.e. $q_i = l_i$ it would follow from part 1 Definition 100 that $Q_A = (l_j(a_i)) = I$ and so P_A would be regular. Thus in general Q_A is regular and $P_X = Q_X Q_A^{-1} P_A$. By definition 4.3 choosing $(\widetilde{q})$ to be the cardinal basis would also imply $Q_X = L_X$ and so $P_X = L_X P_A$.

Part 4 By part 3, $P_X = L_X P_A$ and the regularity of P_A implies $P_X^T \beta = 0$ iff $L_X^T \beta = 0$.

Part 5 That $\dim \operatorname{null} P_X^T = \dim \operatorname{null} L_X^T$ is clear from part 4. Next observe that $\dim \operatorname{null} L_X^T = N - \operatorname{rank} L_X^T = N - \operatorname{rank} L_X$. Now recall the discussion of permutation operators and matrices in Definition 100. Let π permute X so the first M elements of X lie in A. Then $\operatorname{rank} L_X = \operatorname{rank} L_{\pi(X)}$ and since the first M rows of $L_{\pi(X)}$ are $l_i(a_j)$ they compose the unit matrix and we can conclude that $\operatorname{rank} L_{\pi(X)} = M$ and $\dim \operatorname{null} L_X^T = N - M$, as claimed. ■

Theorem 105 ***Properties of unisolvency matrices*** *Suppose $X = \{x^{(i)}\}_{i=1}^{N}$ is a unisolvent set of points in $\mathbb{R}^d$ and $X_1 = \{x^{(i)}\}_{i=1}^{M}$ is a minimal unisolvent subset. Set $X_2 = \{x^{(i)}\}_{i=M+1}^{N}$. Then:*

1. *The cardinal unisolvency matrix w.r.t. X_1 is $L_X = \begin{pmatrix} I_M \\ L_{X_2} \end{pmatrix}$ where $L_{X_2} = \left(l_j\left(x^{(i)}\right)\right)$, $i > M$. Also*
$$P_X^T \beta = 0 \text{ iff } \beta_k = -\sum_{j=M+1}^{N} \beta_j l_k\left(x^{(j)}\right), \quad k = 1, \ldots, M.$$

2. *Suppose* $\beta = \begin{pmatrix} \beta' \\ \beta'' \end{pmatrix}$ *where* $\beta' \in \mathbb{R}^M$, $\beta'' \in \mathbb{R}^{N-M}$. *Then*

$$P_X^T \beta = 0 \text{ iff } \beta' = -L_{X_2}^T \beta'' \text{ iff } \beta = \begin{pmatrix} -L_{X_2}^T \\ I_{N-M} \end{pmatrix} \beta''.$$

3. *Define the* $N \times N$ *matrix* $L_{X;0} = \begin{pmatrix} L_X & O_{N,N-M} \end{pmatrix}$. *Then* $L_{X;0} L_X = L_X$ *and* $\left(L_{X;0}\right)^2 = L_{X;0}$.

Proof. Parts 1 and 2 Since by definition $L_X = \left(l_j\left(x^{(i)}\right)\right)$ and by definition $l_j\left(x^{(i)}\right) = \delta_{i,j}$ when $i \leq M$, we obtain the block form $L_X = \begin{pmatrix} I_M \\ L_{X_2} \end{pmatrix}$. The rest of the theorem follows easily by observing that $L_X^T \beta = \beta' + L_{X_2}^T \beta''$.

Part 3. $L_{X;0} = \begin{pmatrix} L_X & O_{N,M-N} \end{pmatrix} = \begin{pmatrix} I_M & O \\ L_{X_2} & O_{N-M} \end{pmatrix}$, and so

$$L_{X;0} L_X = \begin{pmatrix} I_M & O \\ L_{X_2} & O_{N-M} \end{pmatrix} \begin{pmatrix} I_M \\ L_{X_2} \end{pmatrix} = L_X.$$

Further

$$\left(L_{X;0}\right)^2 = \begin{pmatrix} I_M & O \\ L_{X_2} & O_{N-M} \end{pmatrix} \begin{pmatrix} I_M & O \\ L_{X_2} & O_{N-M} \end{pmatrix} = \begin{pmatrix} I_M & O \\ L_{X_2} & O_{N-M} \end{pmatrix} = L_{X;0}.$$

■

4.3 Properties of the function $\mathcal{Q}_x\left(e^{ix\xi}\right)$

The operators $\mathcal{P}$ and $\mathcal{Q}$ were introduced in the previous section and defined using a minimal unisolvent set of points in $\mathbb{R}^d$. In this section we study the functions $D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)$ and obtain smoothness results and upper bounds, both near the origin and near infinity. Finally, the function $D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)$ is related to the weight function by an upper bound for $\int \frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right|^2}{w|\cdot|^{2\theta}}$. These results will be used in the next section to derive an 'inverse Fourier transform' theorem which expresses the value of a function $f \in X_w^\theta$ in terms of its Fourier transform on $\mathbb{R}^d \backslash 0$.

The next theorem derives some smoothness properties of the function $\mathcal{Q}_x\left(e^{ix\xi}\right)$.

Theorem 106 *Suppose* $\theta \geq 1$ *is an integer and* $M = \dim P_\theta$. *Let* $\{l_i\}_{i=1}^M$ *be the cardinal basis polynomials of* P_θ *associated with the minimal unisolvent set* $A = \{a_i\}_{i=1}^M$. *Then*

$$\mathcal{Q}_x\left(e^{ix\xi}\right) = e^{ix\xi} - \sum_{i=1}^M l_i(x) e^{ia_i\xi}, \quad \xi, x \in \mathbb{R}^d,$$

and $\mathcal{Q}_x\left(e^{ix\xi}\right)$ *has the following local properties.*

1. $\mathcal{Q}_x\left(e^{ix\xi}\right) = 0$ *when* $x \in A$.
2. *If* $|\beta| < \theta$, *then for each* x, $D_\xi^\beta \mathcal{Q}_x\left(e^{ix\xi}\right) \in C_{\emptyset,\theta}^\infty \cap C_B^\infty$.
3. *If* $|\beta| < \theta$, *then for each* x, $D_x^\beta \mathcal{Q}_x\left(e^{ix\xi}\right) \in C_{\emptyset,\theta}^\infty \cap C_{BP}^\infty$.
4. *If* $|\alpha| = \theta$ *and* $|\beta| < \theta$, *then for each* x, $(i\xi)^\alpha D_x^\beta \mathcal{Q}_x\left(e^{ix\xi}\right) \in C_{\emptyset,2\theta}^\infty \cap C_{BP}^\infty$.

Proof. Part 1 Follows directly from the fact that $l_i\left(a_\beta\right) = \delta_{\alpha,\beta}$.

Part 2 Clearly, for each x, $\mathcal{Q}_x\left(e^{-ix\xi}\right) \in C^\infty$. Since

$$D_\xi^\beta \mathcal{Q}_x\left(e^{-ix\xi}\right) = (-ix)^\beta e^{-ix\xi} - \sum_{i=1}^M \left(-ia_i\right)^\beta e^{-ia_i\xi} l_i(x), \tag{4.5}$$

when $\xi = 0$

$$D_\xi^\beta \mathcal{Q}_x \left(e^{-ix\xi}\right) = (-i)^{|\beta|} \left(x^\beta - \sum_{i=1}^M (a_i)^\beta \, l_i(x) \right) = (-i)^{|\beta|} \, \mathcal{Q}\left(x^\beta\right) = 0,$$

when $|\beta| < \theta$. Hence for each x, $D_\xi^\beta \mathcal{Q}_x \left(e^{-ix\xi}\right) \in C_{\emptyset,\theta}^\infty$.

Part 3 Suppose $|\beta| < \theta$.

$$\begin{aligned} D_x^\beta \mathcal{Q}_x \left(e^{-ix\xi}\right) = D_x^\beta \left(e^{-ix\xi} - \sum_{i=1}^M e^{-ia_i\xi} l_i(x) \right) &= (-i\xi)^\beta \, e^{-ix\xi} - \sum_{i=1}^M e^{-ia_i\xi} \left(D^\beta l_i\right)(x) \\ &= e^{-ix\xi} \left((-i\xi)^\beta - \sum_{i=1}^M e^{i(x-a_i)\xi} \left(D^\beta l_i\right)(x) \right). \end{aligned}$$

Again noting Theorem 15, since $e^{-ix\xi}$ is in $C_{\emptyset,0}^\infty$ as a function of ξ, proving that the second factor of the last term is in $C_{\emptyset,\theta}^\infty$ for any fixed x proves the result. In fact, if $|\gamma| < \theta$ and $\gamma \neq \beta$, at $\xi = 0$

$$\begin{aligned} D_\xi^\gamma \left((-i\xi)^\beta - \sum_{i=1}^M e^{i(x-a_i)\xi} \left(D^\beta l_i\right)(x) \right) &= -i^{|\gamma|} \sum_{i=1}^M (x - a_i)^\gamma \left(D^\beta l_i\right)(x) \\ &= -i^{|\gamma|} \left\{ D_x^\beta \mathcal{P}_x \left[(y-x)^\gamma \right] \right\}_{y=x} \\ &= -i^{|\gamma|} \left\{ D_x^\beta \left[(y-x)^\gamma \right] \right\}_{y=x} \\ &= 0, \end{aligned}$$

since $\gamma \neq \beta$. There remains the case of $\gamma = \beta$, but this follows easily using the same technique.

Part 4 Suppose $|\alpha| = \theta$ and $|\beta| < \theta$. First observe that by theorem 15, $(i\xi)^\alpha \in C_{\emptyset,\theta}^\infty$ and since it was shown in part 3 that for each x, $D_x^\beta \mathcal{Q}_x \left(e^{-ix\xi}\right) \in C_{\emptyset,\theta}^\infty$, this part is proved. ■

To prove the next result we will require the following lemma concerning differentiation under the integral sign. This result is Proposition 7.8.4 of Malliavin [14].

Lemma 107 *Suppose $f : \mathbb{R}^{m+n} \to \mathbb{C}$ and we write $f(\xi, x)$ where $\xi \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$. Suppose that:*

1. *For each ξ, $f(\xi, \cdot) \in C^{(k)}(\mathbb{R}^n)$.*
2. *For each x, $\int \left| D_\xi^\alpha f(\xi, x) \right| d\xi < \infty$ for $|\alpha| \leq k$.*

Then when $|\alpha| \leq k$ we have

$$D_x^\alpha \int f(\xi, x) \, d\xi = \int D_x^\alpha f(\xi, x) \, d\xi,$$

and $\int f(\xi, \cdot) \, d\xi \in C^{(k)}(\mathbb{R}^n)$.

In the next lemma we elucidate some of the structure of the function $\mathcal{Q}_x \left(e^{ix\xi}\right)$ studied in the previous theorem.

Lemma 108 *Suppose the operator $\mathcal{Q}$ is generated by the minimal θ-unisolvent set $A = \{a_i\}_{i=1}^M$ and the cardinal basis $\{l_i\}_{i=1}^M$. Then:*

1.
$$\mathcal{Q}_x \left(e^{ix\xi}\right) = \sum_{|\alpha|=\theta} v_\alpha(x, \xi) \frac{(i\xi)^\alpha}{\alpha!}, \tag{4.6}$$

   where
$$\nu_\alpha(x, \xi) = \mathcal{Q}_x \left(x^\alpha \mu_\theta(\xi x) \right),$$

   and
$$\mu_\theta(t) = \frac{e^{it}}{(\theta - 1)!} \int_0^1 e^{-ist} s^{\theta-1} ds, \quad t \in \mathbb{R}.$$

2. $\mu_\theta \in C_B^\infty(\mathbb{R})$ *and* $\left\|D^k\mu_\theta\right\|_\infty < \frac{1}{\theta!}$ *for all* k.

3. For each x, $v_\alpha(x,\cdot) \in C_{BP}^\infty\left(\mathbb{R}^d\right)$ *and for each* ξ, $v_\alpha(\cdot,\xi) \in C_{BP}^\infty\left(\mathbb{R}^d\right)$.
Also, for all $|\gamma| < \theta$

$$\left|D_x^\gamma \nu_\alpha(x,\xi)\right| \leq \left(2^{|\gamma|} + |A|\, c_{A,\gamma}\right)(1+|\xi|)^{|\gamma|}(1+|x|)^\theta, \quad x,\xi \in \mathbb{R}^d,$$

where $|A|$ *and* $c_{A,\gamma}$ *are defined by 4.13 and 4.12 of the proof.*

Proof. Parts 1 and 2 We start by expanding e^{it}, $t \in \mathbb{R}$ about zero using Taylor's theorem with remainder, as given in Appendix A.8, and then substitute $t = x\xi$ to obtain

$$e^{ix\xi} = p_\theta(ix\xi) + \frac{(ix\xi)^\theta e^{ix\xi}}{(\theta-1)!}\int_0^1 e^{-isx\xi} s^{\theta-1} ds, \tag{4.7}$$

where $p_\theta \in P_\theta$. For notational compactness set

$$\mu_\theta(t) = \frac{e^{it}}{(\theta-1)!}\int_0^1 e^{-ist}s^{\theta-1}ds, \quad t \in \mathbb{R}.$$

Next observe that $\mu_\theta \in C_{BP}^\infty$, since μ_θ is e^{it} times the Fourier transform of a bounded L_{loc}^1 function with compact support. Further, since the derivatives of the integrand which defines μ_θ are L^1 functions on $\mathbb{R}$, Lemma 107 enables us to differentiate under the integral sign and this implies that all the derivatives are bounded i.e. $\mu_\theta \in C_B^\infty$. In fact, for any integer $k \geq 0$

$$\begin{aligned} D^k\mu_\theta(t) = D^k\frac{e^{it}}{(\theta-1)!}\int_0^1 e^{-ist}s^{\theta-1}ds &= D^k\frac{1}{(\theta-1)!}\int_0^1 e^{i(1-s)t}s^{\theta-1}ds \\ &= \frac{(i)^k}{(\theta-1)!}\int_0^1 e^{i(1-s)t}(1-s)^k s^{\theta-1}ds, \end{aligned}$$

so that

$$D^k\mu_\theta(0) = \frac{(i)^k}{(\theta-1)!}\frac{k!\,(\theta-1)!}{(k+\theta)!} = \frac{(i)^k\,k!}{(k+\theta)!},$$

and

$$\left|D^k\mu_\theta(t)\right| \leq \frac{1}{(\theta-1)!}\int_0^1 (1-s)^k s^{\theta-1}ds \leq \frac{1}{(\theta-1)!}\int_0^1 s^{\theta-1}ds = \frac{1}{\theta!}.$$

We now rewrite 4.7 as

$$e^{ix\xi} = p_\theta(ix\xi) + (ix\xi)^\theta \mu_\theta(x\xi). \tag{4.8}$$

But for given ξ, $p_\theta(ix\xi)$ is a polynomial in x of degree less than θ, so by part 2 of Theorem 102 $\mathcal{Q}_x(p_\theta(ix\xi)) = 0$ for all x. Thus

$$\begin{aligned} \mathcal{Q}_x\left(e^{ix\xi}\right) &= \mathcal{Q}_x\left((i\xi x)^\theta \mu_\theta(x\xi)\right) \qquad (4.9)\\ &= \mathcal{Q}_x\left(\left(\sum_{|\alpha|=\theta}\frac{x^\alpha(i\xi)^\alpha}{\alpha!}\right)\mu_\theta(x\xi)\right) \\ &= \sum_{|\alpha|=\theta}\frac{1}{\alpha!}\mathcal{Q}_x\left(x^\alpha\mu_\theta(x\xi)\right)(i\xi)^\alpha \\ &= \sum_{|\alpha|=\theta} v_\alpha(x,\xi)(i\xi)^\alpha, \qquad (4.10) \end{aligned}$$

where we have defined the functions $\{v_\alpha\}_{|\alpha|=\theta}$ by

$$v_\alpha(x,\xi) = \frac{1}{\alpha!}\mathcal{Q}_x\left(x^\alpha\mu_\theta(x\xi)\right), \quad |\alpha| = \theta. \tag{4.11}$$

Part 3 Since $\mu_\theta \in C_B^\infty(\mathbb{R})$, for each x, $v_\alpha(x,\cdot) \in C_{BP}^\infty(\mathbb{R}^d)$ and for each ξ, $v_\alpha(\cdot,\xi) \in C_{BP}^\infty(\mathbb{R}^d)$. Now $\{l_i\}$ is the cardinal basis corresponding to the minimal unisolvent set $A = \{a_i\}_{i=1}^M$and there exist constants $c_{A,\gamma}$ such that

$$\sum_{i=1}^M |D^\gamma l_i(x)| \le c_{A,\gamma} (1+|x|)^{\theta-|\gamma|-1}, \quad |\gamma| < \theta,\ x \in \mathbb{R}^d. \tag{4.12}$$

Also define $|A|$ by

$$|A| = \max_k |a_k|. \tag{4.13}$$

We now differentiate equation 4.11 of part 1 with respect to x.

$$\begin{aligned}
(D_x^\gamma v_\alpha)(x,\xi) &= \frac{1}{\alpha!} D_x^\gamma \mathcal{Q}_x (x^\alpha \mu_\theta(x\xi)) \\
&= \frac{1}{\alpha!} D_x^\gamma \left(x^\alpha \mu_\theta(x\xi) - \sum_{i=1}^M (a_i)^\alpha \mu_\theta(a_i\xi) l_i(x) \right) \\
&= \frac{1}{\alpha!} D_x^\gamma (x^\alpha \mu_\theta(x\xi)) - \sum_{i=1}^M (a_i)^\alpha \mu_\theta(a_i\xi) D^\gamma l_i(x) \\
&= \frac{1}{\alpha!} \sum_{\substack{\beta\le\gamma \\ \beta\le\alpha}} \binom{\gamma}{\beta} D^\gamma (x^\alpha) D_x^{\gamma-\beta} (\mu_\theta(x\xi)) - \sum_{i=1}^M (a_i)^\alpha \mu_\theta(a_i\xi) D^\gamma l_i(x) \\
&= \frac{1}{\alpha!} \sum_{\substack{\beta\le\gamma \\ \beta\le\alpha}} \binom{\gamma}{\beta} \binom{\alpha}{\beta} (x^{\alpha-\beta}) \xi^{\gamma-\beta} \left(D^{|\gamma-\beta|} \mu_\theta \right)(x\xi) - \\
&\quad - \frac{1}{\alpha!} \sum_{i=1}^M (a_i)^\alpha \mu_\theta(a_i\xi) D^\gamma l_i(x).
\end{aligned}$$

since $\left|D^k \mu_\theta(t)\right| \le \frac{1}{\theta!}$ for all integers $\theta \ge 1$, we have

$$\begin{aligned}
|(D_x^\gamma v_\alpha)(x,\xi)| &\le \frac{1}{\alpha!} \sum_{\substack{\beta\le\gamma \\ \beta\le\alpha}} \binom{\gamma}{\beta} \binom{\alpha}{\beta} \left|x^{\alpha-\beta}\right| \left|\xi^{\gamma-\beta}\right| + \frac{1}{\alpha!} |A| c_{A,\gamma} (1+|x|)^{\theta-|\gamma|-1} \\
&\le \sum_{\substack{\beta\le\gamma \\ \beta\le\alpha}} \binom{\gamma}{\beta} |x|^{\theta-|\beta|} |\xi|^{|\gamma|-|\beta|} + |A| c_{A,\gamma} (1+|x|)^{\theta-|\gamma|-1} \\
&\le \sum_{\beta\le\gamma} \binom{\gamma}{\beta} (1+|x|)^\theta (1+|\xi|)^{|\gamma|} + |A| c_{A,\gamma} (1+|x|)^{\theta-|\gamma|-1} \\
&= 2^{|\gamma|} (1+|x|)^\theta (1+|\xi|)^{|\gamma|} + |A| c_{A,\gamma} (1+|x|)^{\theta-|\gamma|-1} \\
&\le 2^{|\gamma|} (1+|x|)^\theta (1+|\xi|)^{|\gamma|} + |A| c_{A,\gamma} (1+|x|)^\theta \\
&\le \left(2^{|\gamma|} + |A| c_{A,\gamma}\right) (1+|x|)^\theta (1+|\xi|)^{|\gamma|}.
\end{aligned}$$

■

Theorem 109 *Suppose the operator $\mathcal{Q}$ is generated by the minimal θ-unisolvent set $A = \{a_i\}_{i=1}^M$ and the cardinal basis $\{l_i\}_{i=1}^M$. Then for $x, \xi \in \mathbb{R}^d$:*

1. *For all γ we have the simple but useful estimate*

$$\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| \le |\xi|^{|\gamma|} + \sum_{k=1}^M |(D^\gamma l_k)(x)|.$$

2. *For all γ*

$$\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| \le \begin{cases} k_{\gamma,A,\theta} (1+|x|)^\theta |\xi|^\theta (1+|\xi|)^{|\gamma|}, & |\gamma| < \theta, \\ |\xi|^{|\gamma|}, & |\gamma| \ge \theta. \end{cases}$$

This inequality is useful for ξ near the origin.

3. For all γ

$$\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| \leq \begin{cases} k'_{\gamma,A}\left(1+|x|\right)^{\theta-1-|\gamma|}\left(1+|\xi|\right)^{|\gamma|}, & |\gamma|<\theta, \\ |\xi|^{|\gamma|}, & |\gamma|\geq\theta. \end{cases}$$

This inequality is useful for all x *or for when* ξ *is large, as the estimates have lower powers than those in part 2.*

The constants $k_{\gamma,A,\theta}$ *and* $k'_{\gamma,A}$ *are defined in the proof.*

Proof. Part 1 By definition of $\mathcal{Q}_x$

$$D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right) = \left(i\xi\right)^\gamma e^{ix\xi} - \sum_{k=1}^{M}\left(D^\gamma l_k\right)(x) e^{ia_k\xi},$$

so it follows that

$$\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| \leq |\xi|^{|\gamma|} + \sum_{k=1}^{M}\left|\left(D^\gamma l_k\right)(x)\right|.$$

Part 2 Since $|\gamma|<\theta$, we can apply the estimate of part 3 of Lemma 108 to the formula 4.8 for $\mathcal{Q}_x\left(e^{ix\xi}\right)$ of part 1 of Lemma 108.

$$\begin{aligned}\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| = \left|\sum_{|\alpha|=\theta} D_x^\gamma v_\alpha\left(x,\xi\right)\left(i\xi\right)^\alpha\right| &\leq |\xi|^\theta \sum_{|\alpha|=\theta}\left|D_x^\gamma v_\alpha\left(x,\xi\right)\right| \\ &\leq |\xi|^\theta \sum_{|\alpha|=\theta}\left(2^{|\gamma|}+|A|\,c_{A,\gamma}\right)\left(1+|x|\right)^\theta\left(1+|\xi|\right)^{|\gamma|} \\ &< \left(\sum_{|\alpha|=\theta}1\right)\left(2^{|\gamma|}+|A|\,c_{A,\gamma}\right)\left(1+|x|\right)^\theta|\xi|^\theta\left(1+|\xi|\right)^{|\gamma|} \\ &= k_{\gamma,A,\theta}\left(1+|x|\right)^\theta|\xi|^\theta\left(1+|\xi|\right)^{|\gamma|},\end{aligned}$$

where $k_{\gamma,A,\theta} = \left(\sum_{|\alpha|=\theta}1\right)\left(2^{|\gamma|}+|A|\,c_{A,\gamma}\right)$ and $|A|$ and $c_{A,\gamma}$ are given by 4.13 and 4.12.

On the other hand, if $|\gamma|\geq\theta$, $\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| = \left|\xi^\gamma\right| \leq |\xi|^{|\gamma|}$.

Part 3 From part 1, if $|\gamma|<\theta$

$$\begin{aligned}\left|D_x^\gamma \mathcal{Q}_x\left(e^{ix\xi}\right)\right| \leq |\xi|^{|\gamma|} + \sum_{k=1}^{M}\left|\left(D^\gamma l_k\right)(x)\right| &\leq \left(1+|\xi|\right)^{|\gamma|} + c_{A,\gamma}\left(1+|x|\right)^{\theta-1-|\gamma|} \\ &\leq 2\max\left\{1,c_{A,\gamma}\right\}\left(1+|\xi|\right)^{|\gamma|}\left(1+|x|\right)^{\theta-1-|\gamma|} \\ &= k'_{\gamma,A}\left(1+|\xi|\right)^{|\gamma|}\left(1+|x|\right)^{\theta-1-|\gamma|},\end{aligned}$$

where $k'_{\gamma,A} = 2\max\left\{1,c_{A,\gamma}\right\}$.

On the other hand, if $|\gamma|\geq\theta$, we use the result of part 2. ■

The next theorem uses the previous theorem to relate the function $D_x^\gamma \mathcal{Q}_x\left(e^{-i(x,\xi)}\right)$, discussed above, to a weight function of order θ.

Theorem 110 *Suppose that the weight function* w *has properties W2.1 and W3 for order* θ *and* κ*, and that the operator* $\mathcal{Q}$ *is defined using a minimal unisolvent set* A *of order* θ*. Define* r_4 *by* $r_4=0$ *if* w *has property W3.1 and* $r_4=r_3$ *if* w *has property W3.2. Then for each* $x\in\mathbb{R}^d$, $\frac{D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)}{\sqrt{w}|\cdot|^{2\theta}}\in L^2$ *whenever* $|\gamma|\leq\kappa$*. In fact, there exists a positive constant* c_w*, independent of* x*, such that*

$$\int\frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right|^2}{w\,|\cdot|^{2\theta}} \leq \begin{cases}\left(c_w\right)^2\left(1+|x|\right)^{2\theta}, & |\gamma|<\theta, \\ \left(c_w\right)^2, & |\gamma|\geq\theta,\end{cases} \tag{4.14}$$

where c_w *is given by 4.20.*

Proof. We want to show that for each x, $\frac{D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\xi)}\right)}{\sqrt{w}|\cdot|^\theta} \in L^2$ whenever $|\gamma| \le \kappa$. We start by splitting the domain of integration in the manner used to define the function ζ. There are two cases: $|\gamma| < \theta$ and $|\gamma| \ge \theta$. Set $m = \min\{\theta - 1, \kappa\}$.

$\boxed{\textbf{Case 1}\ |\gamma| < \theta,\ |\gamma| \le \kappa}$ Part 2 of Theorem 109 implies

$$
\begin{aligned}
\int\limits_{|\cdot| \le r_4} \frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right|^2}{w\,|\cdot|^{2\theta}} &\le \int\limits_{|\cdot| \le r_4} \frac{\left(k_{\gamma,A,\theta}\,|\cdot|^\theta\,(1+|\cdot|)^{|\gamma|}\,(1+|x|)^\theta\right)^2}{w\,|\cdot|^{2\theta}} \\
&\le (1+r_4)^{2|\gamma|}\,(k_{\gamma,A,\theta})^2 \left(\int\limits_{|\cdot| \le r_4} \frac{1}{w}\right)(1+|x|)^{2\theta} \\
&\le (1+r_4)^{2\kappa} \left(\max_{|\gamma|<\theta} k_{\gamma,A,\theta}\right)^2 \left(\int\limits_{|\cdot| \le r_4} \frac{1}{w}\right)(1+|x|)^{2\theta}, \qquad (4.15)
\end{aligned}
$$

which exists since property W2.1 is $1/w \in L^1_{loc}$. Next we use part 3 of Theorem 109.

$$
\begin{aligned}
\int\limits_{|\cdot| \ge r_4} \frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right|^2}{w\,|\cdot|^{2\theta}} &\le \int\limits_{|\cdot| \ge r_4} \frac{\left(k'_{\gamma,A}\,(1+|\cdot|)^{|\gamma|}\,(1+|x|)^{\theta-1-|\gamma|}\right)^2}{w\,|\cdot|^{2\theta}} \\
&= \left(k'_{\gamma,A}\right)^2 \left(\int\limits_{|\cdot| \ge r_4} \frac{(1+|\cdot|)^{2|\gamma|}}{w\,|\cdot|^{2\theta}}\right)(1+|x|)^{2(\theta-1-|\gamma|)} \\
&\le 2^{2|\gamma|}\left(k'_{\gamma,A}\right)^2 \left(\int\limits_{|\cdot| \ge r_4} \frac{1+|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}}\right)(1+|x|)^{2\theta} \\
&\le 2^{2\kappa}\left(\max_{|\gamma|<\theta} k'_{\gamma,A}\right)^2 \max_{|\gamma|\le\kappa}\left(\int\limits_{|\cdot| \ge r_4} \frac{1+|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}}\right)(1+|x|)^{2\theta}, \qquad (4.16)
\end{aligned}
$$

which exists by part 2 of Theorem 11. Combining 4.15 and 4.16 we have the estimate

$$
\int \frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\xi)}\right)\right|^2}{w\,|\cdot|^{2\theta}} \le \left(c'_w\right)^2 (1+|x|)^{2\theta}, \qquad (4.17)
$$

where

$$
c'_w = \max\left\{2^\kappa \max_{|\gamma|<\theta} k'_{\gamma,A},\ (1+r_4)^\kappa \max_{|\gamma|<\theta} k_{\gamma,A,\theta}\right\} \max_{|\gamma|\le\kappa}\left(\int\limits_{|\cdot| \le r_4} \frac{1}{w} + \int\limits_{|\cdot| \ge r_4} \frac{1+|\cdot|^{2|\gamma|}}{w\,|\cdot|^{2\theta}}\right)^{\frac{1}{2}}. \qquad (4.18)
$$

$\boxed{\textbf{Case 2}\ |\gamma| \geq \theta,\ |\gamma| \leq \kappa}$ Using part 2 of Theorem 109 we obtain

$$\begin{aligned}
\int \frac{\left|D_x^\gamma \mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right|^2}{w\left|\cdot\right|^{2\theta}} &\leq \int\limits_{|\cdot|\leq r_4} \frac{\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} + \int\limits_{|\cdot|\geq r_4} \frac{\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \\
&= \int\limits_{|\cdot|\leq r_4} \frac{\left|\cdot\right|^{2|\gamma|-2\theta}}{w} + \int\limits_{|\cdot|\geq r_4} \frac{\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \\
&\leq (r_4)^{2|\gamma|-2\theta} \int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \\
&\leq (1+r_4)^{2|\gamma|-2\theta} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \right) \\
&\leq (1+r_4)^{2\kappa} \max_{|\gamma|\leq\kappa} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{1+\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \right),
\end{aligned} \tag{4.19}$$

since $|\gamma| \leq \kappa$. We can now combine the results of the two cases by defining the constant c_w to be

$$c_w = \max\left\{ 2^\kappa \max_{|\gamma|\leq\theta} k'_{\gamma,A}, (1+r_4)^\kappa \max_{|\gamma|\leq\theta} k_{\gamma,A,\theta}, (1+r_4)^\kappa \right\} \max_{|\gamma|\leq\kappa} \left(\int\limits_{|\cdot|\leq r_4} \frac{1}{w} + \int\limits_{|\cdot|\geq r_4} \frac{1+\left|\cdot\right|^{2|\gamma|}}{w\left|\cdot\right|^{2\theta}} \right), \tag{4.20}$$

and so from 4.17, 4.18 and 4.19 obtain the estimates 4.14 of this lemma. The constants $k'_{\gamma,A}$ and $k_{\gamma,A,\theta}$ are defined in Theorem 109. ■

4.4 An inverse Fourier transform result for functions in X_w^θ

In this section we use the properties of the function $\mathcal{Q}_x\left(e^{i(x,\xi)}\right)$ derived in the previous section to derive an inverse Fourier transform theorem which expresses the value of a function $f \in X_w^\theta$ in terms of its distribution Fourier transform which is a function in $L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$.

Summary 111 will now recall some results about the semi-Hilbert data space

$$X_w^\theta = \left\{ f \in S' : \widehat{D^\alpha f} \in L^1_{loc}\left(\mathbb{R}^d\right),\ \int w\left|\widehat{D^\alpha f}\right|^2 < \infty \ for\ all\ |\alpha| = \theta \right\},$$

with semi-inner product and seminorm

$$\langle f, g\rangle_{w,\theta} = \sum_{|\alpha|=\theta} \frac{\theta!}{\alpha!} \int w\widehat{D^\alpha f}\,\overline{\widehat{D^\alpha g}}, \quad |f|_{w,\theta} = \sqrt{\langle f, f\rangle_{w,\theta}},$$

introduced as Definition 21.

Summary 111 *Suppose the weight function w has property **W2**. If $f \in X_w^\theta$ then $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and we can define the function $f_F : \mathbb{R}^d \to \mathbb{C}$ a.e. by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Further:*

1. *The seminorm $\left|\cdot\right|_{w,\theta}$ satisfies (part 1 Theorem 25)*

$$\int w\left|\cdot\right|^{2\theta}\left|f_F\right|^2 = \left|f\right|^2_{w,\theta}. \tag{4.21}$$

2. *The functional $\left|\cdot\right|_{w,\theta}$ is a seminorm. In fact,* null $\left|\cdot\right|_{w,\theta} = P_\theta$ *and $X_w^\theta \cap P = P_\theta$ (part 3 Theorem 25).*

3. *$f_F = 0$ iff $f \in P_\theta$ (part 1 Theorem 27).*

4. *$f_F \in S'_{\emptyset,\theta}$ and $\widehat{f} = f_F$ on $S_{\emptyset,\theta}$ (part 2 Theorem 27).*

5. *Since the weight function has property W2 we can use the definition of X_w^θ from Corollary 29, namely*

$$X_w^\theta = \left\{ f \in S' : \widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right), \ \int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 < \infty, \ |\alpha| = \theta \text{ implies } \xi^\alpha \widehat{f} = \xi^\alpha f_F \text{ on } S \right\}. \tag{4.22}$$

Note that this definition makes sense because the conditions $\widehat{f} \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ and $\int w \left|\cdot\right|^{2\theta} \left|f_F\right|^2 < \infty$ imply (part 2 Corollary 24) that when $|\alpha| = \theta$, $\xi^\alpha f_F$ is a regular tempered distribution in the sense of part 2 Appendix A.5.1.

6. *X_w^θ is complete in the seminorm sense (Theorem 37).*

7. *If w also has property W3 for order θ and smoothness parameter κ then $X_w^\theta \subset C_{BP}^{(\lfloor \kappa \rfloor)}$ (Theorem 40).*

Lemma 112 *Suppose the weight function w also has property W2.*
Then for all $f \in X_w^\theta$ and all multi-indexes $\gamma \geq 0$

$$\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) \widehat{f} = \left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) f_F \in S',$$

where the function $f_F : \mathbb{R}^d \to \mathbb{C}$ is defined a.e. by $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$.

Proof. If $|\gamma| < \theta$ then from part 2 Theorem 106 we have $D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right) \in C_{\emptyset,\theta}^\infty \cap C_{BP}^\infty$.
If $|\gamma| \geq \theta$ then $D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\xi)}\right) = (i\xi)^\gamma e^{i(x,\cdot)}$ and by part 3 Theorem 15, $D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\xi)}\right) \in C_{\emptyset,\theta}^\infty \cap C_{BP}^\infty$ for each x. Again by Theorem 15, $\phi D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right) \in S_{\emptyset,\theta}$ if $\phi \in S$. In addition, from part 4 Summary 111, $f_F \in S'_{\emptyset,\theta}$ and $\widehat{f} = f_F$ on $S_{\emptyset,\theta}$. So for $\phi \in S$ we now have

$$\begin{aligned} \left[\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) \widehat{f}, \phi\right] = \left[\widehat{f}, \phi D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right] &= \left[f_F, \phi D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right] \\ &= \left[\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) f_F, \phi\right], \end{aligned}$$

which proves this lemma. ■
The next theorem is an inverse Fourier transform result for functions in X_w^θ.

Theorem 113 *If the weight function w has properties W2.1 and W3 for order $\theta \geq 1$ and $\kappa \geq 0$ then $X_w^\theta \subset C_{BP}^{(\lfloor \kappa \rfloor)}$.*
Further, if $f \in X_w^\theta$, $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$, $x \in \mathbb{R}^d$ and $|\gamma| \leq \lfloor \kappa \rfloor$ then $\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) f_F \in L^1$ and

$$D^\gamma f(x) = (2\pi)^{-\frac{d}{2}} \int \left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\xi)}\right)\right) f_F(\xi)\, d\xi + \left(D^\gamma \mathcal{P} f\right)(x), \tag{4.23}$$

and we have the growth estimates

$$\left|D^\gamma f(x)\right| \leq \begin{cases} c_w \left|f\right|_{w,\theta} (1 + |x|)^\theta + \left|\left(D^\gamma \mathcal{P} f\right)(x)\right|, & |\gamma| < \theta, \\ c_w \left|f\right|_{w,\theta}, & |\gamma| \geq \theta, \end{cases} \tag{4.24}$$

where c_w is defined in Theorem 110.

Proof. First note that $f_F \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$ implies $\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) f_F \in L^1_{loc}\left(\mathbb{R}^d \backslash 0\right)$. For each x, we apply the Cauchy-Schwartz inequality to obtain

$$\begin{aligned} \int \left|\left(D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right) f_F\right| &= \int \frac{\left|D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right|}{\sqrt{w} \left|\cdot\right|^\theta} \sqrt{w} \left|\cdot\right|^\theta \left|f_F\right| \\ &\leq \left(\int \frac{\left|D_x^\gamma \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)\right|^2}{w \left|\cdot\right|^{2\theta}}\right)^{\frac{1}{2}} \left|f\right|_{w,\theta}, \end{aligned} \tag{4.25}$$

and the last integral exists by inequality 4.14 of Theorem 110. From Lemma 112, for each x,

$$\left(D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right) f_F = \left(D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\right)\widehat{f} = \left((i\xi)^{\gamma} e^{i(x,\cdot)} - \sum_{i=1}^{M} e^{i\langle a_i,\cdot\rangle} D^{\gamma}l_i(x)\right)\widehat{f}$$
$$= \left(D^{\gamma}f(\cdot + x) - \sum_{i=1}^{M} f(\cdot + a_i) D^{\gamma}l_i(x)\right)^{\wedge}.$$

Now set $g_x = D^{\gamma}f(\cdot + x) - \sum\limits_{i=1}^{M} D^{\gamma}l_i(x) f(\cdot + a_i)$ so that for each x, $g_x \in S'$ and $\widehat{g_x} = D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\cdot)}\right)\widehat{f} \in L^1$. But by Lemma 48 Chapter 1, if $u \in S'$ and $\widehat{u} \in L^1$, then $u \in C_B^{(0)}$ and $u = (2\pi)^{-d/2}\int e^{i(\cdot,\xi)}\widehat{u}(\xi)\,d\xi$. Thus, $D^{\gamma}f(\cdot + x) - \sum\limits_{i=1}^{M} f(\cdot + a_i) D^{\gamma}l_i(x) \in C_B^{(0)}$ and

$$D^{\gamma}f(\cdot + x) - \sum_{i=1}^{M} f(\cdot + a_i) D^{\gamma}l_i(x) = (2\pi)^{-\frac{d}{2}} \int e^{i(\cdot,\xi)} D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\xi)}\right) f_F(\xi)\,d\xi,$$

which implies

$$D^{\gamma}f(x) = (2\pi)^{-\frac{d}{2}} \int D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\xi)}\right) f_F(\xi)\,d\xi + D^{\gamma}\mathcal{P}f(x),$$

which proves 4.23. Now we use the Cauchy-Schwartz inequality to get

$$\begin{aligned}
|D^{\gamma}f(x)| &\le (2\pi)^{-\frac{d}{2}} \int \left|D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\xi)}\right) f_F(\xi)\right| d\xi + |D^{\gamma}\mathcal{P}f(x)| \\
&\le (2\pi)^{-\frac{d}{2}} \int \frac{\left|D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\xi)}\right)\right|}{\sqrt{w}\,|\cdot|^{\theta}} \sqrt{w}\,|\cdot|^{\theta}\,|f_F(\xi)|\,d\xi + |D^{\gamma}\mathcal{P}f(x)| \\
&\le (2\pi)^{-\frac{d}{2}} \left(\int \frac{\left|D_x^{\gamma}\mathcal{Q}_x\left(e^{i(x,\xi)}\right)\right|^2}{w\,|\cdot|^{2\theta}}\right)^{\frac{d}{2}} |f|_{w,\theta} + |D^{\gamma}\mathcal{P}f(x)|.
\end{aligned}$$

Substituting inequality 4.14 of Theorem 110 yields

$$|D^{\gamma}f(x)| \le \begin{cases} c_w\,|f|_{w,\theta}\,(1+|x|)^{\theta} + |D^{\gamma}\mathcal{P}f(x)|, & |\gamma| < \theta, \\ c_w\,|f|_{w,\theta} & |\gamma| \ge \theta, \end{cases}$$

which proves inequality 4.24. ■

4.5 Riesz representers for the functionals $u \to D^{\gamma}u(x)$

Suppose that the weight function w has properties W2 and W3 for order θ and κ. Then the functions in X_w^{θ} are continuous. In this section we will derive simple explicit formulas for the Riesz representers of the evaluation functionals $f \to D^{\gamma}f(x)$ where $f \in X_w^{\theta}$ and $|\gamma| \le \kappa$ and X_w^{θ} is endowed with a special norm called the Light norm. When $\gamma = 0$ the representer is denoted by R_x and $D_x^{\gamma}R_x$ is shown to be the Riesz representer for the evaluation functional $D^{\gamma}u(x)$ when $|\gamma| \le \lfloor\kappa\rfloor$. The formula for R_x is exhibited in equation 4.35 and is expressed in terms of a basis function and the cardinal polynomial basis used to define the norm. The existence of $R_x \in X_w^{\theta}$ means that X_w^{θ} is a reproducing kernel Hilbert space of continuous functions.

The approach to deriving the Riesz representers is to first introduce a special subspace of X_w^{θ} in the next subsection and then to deduce the form R_x must have from the definition of R_x and its implications for this subspace of functions. The operators $\mathcal{P}$ and $\mathcal{Q}$ introduced in Section 4.2.1 are also useful.

4.5.1 The functions $G * \widehat{S}_{\emptyset,\theta}$

We will now prove some properties of the space of functions, denoted by $G * \widehat{S}_{\emptyset,\theta}$, having the form $G * \varphi$ where $\varphi \in \widehat{S}_{\emptyset,\theta}$. Here $\widehat{S}_{\emptyset,\theta}$ denotes the Fourier transforms of the functions in $S_{\emptyset,\theta}$. In the next section these functions will help us deduce the precise form of the Riesz representer of the functionals $f \to D^{\alpha}f(x)$.

Corollary 114 *Suppose the weight function w has property W2, and suppose G is a basis distribution of order $\theta \geq 1$ generated by w. Then $\varphi \in \widehat{S}_{\emptyset,\theta}$ implies $G * \varphi \in X_w^\theta$ and*

$$|G * \varphi|_{w,\theta}^2 = \int \frac{|\widehat{\varphi}|^2}{w\,|\cdot|^{2\theta}} = \left[\widehat{G}, |\widehat{\varphi}|^2\right]. \tag{4.26}$$

Also

$$\langle f, G * \varphi \rangle_{w,\theta} = [f, \overline{\varphi}], \quad f \in X_w^\theta. \tag{4.27}$$

Proof. The first step is to show that $G * \varphi \subset X_w^\theta$. Let $g = G * \varphi$ where $\varphi \in \widehat{S}_{\emptyset,\theta}$. To prove that $g \in X_w^\theta$ we use definition 4.22 of X_w^θ i.e. we show that $g \in S'$, $\widehat{g} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$, $\int w\,|\cdot|^{2\theta}\,|g_F|^2 < \infty$ and $\xi^\alpha \widehat{g} = \xi^\alpha g_F$ on S.

Since $G \in S'$ and $\varphi \in S$, $g = G * \varphi \in S'$ and $\widehat{g} = \widehat{\varphi}\,\widehat{G}$, as distributions. Further, from Definition 44 of the basis distribution G we know that $\left[\widehat{G}, \psi\right] = \int \frac{\psi}{w|\cdot|^{2\theta}}$ for $\psi \in S_{\emptyset,2\theta}$, and so $[\widehat{g}, \psi] = \int \frac{\widehat{\varphi}\psi}{w|\cdot|^{2\theta}}$ for $\psi \in S_{\emptyset,2\theta}$. But property W2.1 implies $\frac{\widehat{\varphi}}{w|\cdot|^{2\theta}} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ which means that $g_F = \frac{\widehat{\varphi}}{w|\cdot|^{2\theta}}$ on $\mathbb{R}^d \backslash 0$. But by definition $\widehat{\varphi} \in S_{\emptyset,\theta}$ so that Theorem 15 implies $|\widehat{\varphi}|^2 = \widehat{\varphi}\overline{\widehat{\varphi}} \in S_{\emptyset,2\theta}$ and

$$\int w\,|\cdot|^{2\theta}\,|g_F|^2 = \int w\,|\cdot|^{2\theta}\,\frac{|\widehat{\varphi}|^2}{w^2\,|\cdot|^{4\theta}} = \int \frac{|\widehat{\varphi}|^2}{w\,|\cdot|^{2\theta}} = \left[\widehat{G}, |\widehat{\varphi}|^2\right] < \infty. \tag{4.28}$$

To show $\xi^\alpha \widehat{g} = \xi^\alpha g_F$ on S, we need to show that $\xi^\alpha \widehat{\varphi}\widehat{G} = \frac{\xi^\alpha \widehat{\varphi}}{w|\cdot|^{2\theta}}$ on S. But by Theorem 15, $\xi^\alpha \in C_{\emptyset,\theta}^\infty \cap C_{BP}^\infty$ and hence $\xi^\alpha \widehat{\varphi} \in S_{\emptyset,2\theta}$ and hence for $\psi \in S$ the basis distribution definition implies $\left[\xi^\alpha \widehat{\varphi}\widehat{G}, \psi\right] = \left[\widehat{G}, \xi^\alpha \widehat{\varphi}\psi\right] = \int \frac{\xi^\alpha \widehat{\varphi}\psi}{w|\cdot|^{2\theta}}$, as required. This means that 4.28 is true and so proves 4.26.

Now to prove formula 4.27. From above we know that $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$ then

$$\langle f, G * \varphi \rangle_w = \int w\,|\cdot|^{2\theta}\,f_F \overline{g_F} = \int w\,|\cdot|^{2\theta}\,f_F \frac{\overline{\widehat{\varphi}}}{w\,|\cdot|^{2\theta}} = \int f_F \overline{\widehat{\varphi}} = \int f_F \widehat{\varphi^*},$$

where $\varphi^*(x) = \overline{\varphi(-x)}$ and $\widehat{\phi^*} \in S_{\emptyset,\theta}$. From part 4 of Summary 111 $f_F \in S_{\emptyset,\theta}'$ and $\widehat{f} = f_F$ on $S_{\emptyset,\theta}$, so that

$$\langle f, G * \varphi \rangle_w = \left[\widehat{f}, \widehat{\varphi^*}\right] = [f, \overline{\varphi}],$$

as required. ■

4.5.2 The Light norm for X_w^θ

Using the results of the previous section, a unisolvent set of points will be used to create an inner product space from the semi-inner product space X_w^θ introduced above in Definition 21. This particular inner product was introduced by Light and Wayne in [11] and involves evaluating functions in X_w^θ on the unisolvent set, which is OK since $X_w^\theta \subset C_{BP}^{(\lfloor \kappa \rfloor)}$ by part 5 of Summary 111.

Definition 115 *Suppose $\theta \geq 1$ and suppose the set $A = \{a_i\}$ is minimally unisolvent with respect to the polynomial space P_θ. Then following Light and Wayne [11] we will introduce the following norm and inner product for the space X_w^θ:*

$$(u, v)_{w,\theta} = \langle u, v \rangle_{w,\theta} + \sum_i u(a_i)\,\overline{v(a_i)}, \quad \|u\|_{w,\theta}^2 = |u|_{w,\theta}^2 + \sum_i |u(a_i)|^2, \tag{4.29}$$

where $|u|_{w,\theta}$ is defined by 4.56. $\|\cdot\|_{w,\theta}$ is actually a norm since $\|\cdot\|_{w,\theta} = 0$ implies that $|u|_{w,\theta} = 0$ and $u(a_i) = 0$ for all unisolvent a_i. By part 3 of Summary 111, $|u|_{w,\theta} = 0$ means $u \in P_\theta$ and so from the definition of unisolvent sets $u = 0$.

Theorem 116 *Properties of the Light norm.*

Suppose the Light norm and the Lagrangian interpolation operator $\mathcal{P}$ are defined using the same minimal unisolvent set $\{a_i\}$. Then:

1. $(\mathcal{Q}u, v)_{w,\theta} = (u, \mathcal{Q}v)_{w,\theta} = \langle u, v \rangle_{w,\theta}$.

2. $(\mathcal{P}u, v)_{w,\theta} = (u, \mathcal{P}v)_{w,\theta} = \sum_i u(a_i)\overline{v(a_i)}$.

Proof. Define the Light norm using the minimal unisolvent set $A = \{a_i\}$. Then using results from Theorem 102 concerning the operators $\mathcal{P}$ and $\mathcal{Q}$:

Part 1 $(\mathcal{Q}u, v)_{w,\theta} = \langle \mathcal{Q}u, v\rangle_{w,\theta} + \sum_i \mathcal{Q}u(a_i)\overline{v(a_i)} = \langle \mathcal{Q}u, v\rangle_{w,\theta}$, since $\mathcal{Q}u(a_i) = 0$. But $\langle \mathcal{Q}u, v\rangle_{w,\theta} = \langle u - \mathcal{P}u, v\rangle_{w,\theta} = \langle u, v\rangle_{w,\theta}$ and so $(\mathcal{Q}u, v)_{w,\theta} = \langle u, v\rangle_{w,\theta}$. Similarly $(u, \mathcal{Q}v)_{w,\theta} = \langle u, v\rangle_{w,\theta}$.

Part 2 $(\mathcal{P}u, v)_{w,\theta} = \langle \mathcal{P}u, v\rangle_{w,\theta} + \sum_i \mathcal{P}u(a_i)\overline{v(a_i)} = \sum_i \mathcal{P}u(a_i)\overline{v(a_i)} =$
$= \sum_i u(a_i)\overline{v(a_i)}$, since $\mathcal{P}u$ interpolates $\{a_i, u(a_i)\}$. Similarly $(u, \mathcal{P}v)_{w,\theta} = \sum_i u(a_i)\overline{v(a_i)}$. ■

4.5.3 Derivation of the Riesz representers $D_x^\gamma R_x$

In this subsection we derive simple explicit formulas for the Riesz representers for the evaluation functionals where $D^\gamma u \in X_w^\theta$ and $|\gamma| \le \lfloor\kappa\rfloor$. These formulas are exhibited in equation 4.35 below and involve order γ derivatives of the basis function and the cardinal basis functions associated with the Light norm.

Suppose that the weight function w has properties W2 and W3 for order θ and parameter κ. Then by Summary 111 the functions in X_w^θ have order $\lfloor\kappa\rfloor$ differentiability. We endow the space X_w^θ with the Light norm generated by a minimal unisolvent set $A = \{a_i\}_{i=1}^M$ and its corresponding cardinal basis $\{l_i\}_{i=1}^M$ of P_θ. We will also make good use of the Lagrangian interpolation operator $\mathcal{P}$ and the operator $\mathcal{Q} = I - \mathcal{P}$ introduced in Section 4.2.1.

The approach used here is to first assume the Riesz representers exist, and then to derive the form of the representer by substituting for u two classes of functions. First we use the cardinal basis functions associated with the unisolvent set used to define the Light norm and then the functions in $G * \widehat{S}_{\emptyset,\theta}$ discussed in Subsection 4.5.1. By this means we obtain the form of the representers. Once we have the candidates for Riesz representers it will be shown that they have the properties initially assumed.

We start by assuming that for all $x \in \mathbb{R}^d$ there exists $R_x \in X_w^\theta$ such that $(u, R_x)_{w,\theta} = u(x)$ for all $u \in X_w^\theta$. Now if $h \in \mathbb{R}^d$ then $u(x+h) - u(x) = (u, R_{x+h})_{w,\theta} - (u, R_x)_{w,\theta} = (u, R_{x+h} - R_x)_{w,\theta}$ and so we would expect the formula $D_k u(x) = \left(u, \frac{\partial R_x}{\partial x_k}\right)_{w,\theta}$ to hold, and indeed we would expect the formulas

$$D^\gamma u(x) = (u, D_x^\gamma R_x)_{w,\theta}, \quad |\gamma| \le \lfloor\kappa\rfloor, \tag{4.30}$$

to hold. Next observe that we have the anti-symmetric result

$$R_x(y) = (R_x, R_y)_{w,\theta} = \overline{(R_y, R_x)_{w,\theta}} = \overline{R_y(x)}. \tag{4.31}$$

By definition the cardinal basis $\{l_i\}_{i=1}^M$ of P_θ has real-valued coefficients and satisfies $l_i = (a_j)\delta_{i,j}$. Thus

$$l_j(x) = (l_j, R_x)_{w,\theta} = \langle l_j, R_x\rangle_{w,\theta} + \sum_{i=1}^M l_j(a_i)\overline{R_x(a_i)} = \overline{R_x(a_j)} = R_x(a_j),$$

so that

$$l_j(x) = R_x(a_j). \tag{4.32}$$

Next we will substitute functions from the space $G * \widehat{S}_{\emptyset,\theta} = \left\{G * \phi : \phi \in \widehat{S}_{\emptyset,\theta}\right\}$ introduced in Section 4.5.1. Thus, if $\phi \in \widehat{S}_{\emptyset,\theta}$ then by Corollary 114 $G * \phi \in X_w^\theta$ and so

$$\begin{aligned}
\overline{(G*\phi)(x)} = \overline{(G*\phi, R_x)_{w,\theta}} &= (R_x, G*\phi)_{w,\theta} \\
&= \langle R_x, G*\phi\rangle_{w,\theta} + \sum_{i=1}^M R_x(a_i)\overline{(G*\phi)(a_i)} \\
&= \langle R_x, G*\phi\rangle_{w,\theta} + \sum_{i=1}^M l_i(x)\overline{(G*\phi)(a_i)} \\
&= \left[R_x, \overline{\phi}\right] + \sum_{i=1}^M l_i(x)\overline{(G*\phi)(a_i)},
\end{aligned}$$

where the last line was derived using equation 4.27. Rearrange this equation to get

$$\left[R_x,\overline{\phi}\right]=\overline{(G*\phi)(x)}-\sum_{i=1}^{M}l_i(x)\,\overline{(G*\phi)(a_i)}.$$

But

$$\begin{aligned}\overline{(G*\phi)(x)}=\overline{(G*\phi)(x)}=(2\pi)^{-\frac{d}{2}}\,\overline{[G(x-\cdot),\phi]}&=(2\pi)^{-\frac{d}{2}}\left[\overline{G(x-\cdot)},\overline{\phi}\right]\\&=(2\pi)^{-\frac{d}{2}}\left[G(\cdot-x),\overline{\phi}\right],\end{aligned}$$

so

$$\left[R_x,\overline{\phi}\right]=(2\pi)^{-\frac{d}{2}}\left[G(\cdot-x),\overline{\phi}\right]-(2\pi)^{-\frac{d}{2}}\sum_{i=1}^{M}l_i(x)\left[G(\cdot-a_i),\overline{\phi}\right],$$

or

$$\left[R_x,\overline{\phi}\right]=(2\pi)^{-\frac{d}{2}}\left[G(\cdot-x)-\sum_{i=1}^{M}l_i(x)\,G(\cdot-a_i),\overline{\phi}\right],$$

for all $\phi\in\widehat{S}_{\emptyset,\theta}$. But by part 1 of Theorem 17, $u\in\widehat{P}_n$ iff $[u,\phi]=0$ for all $\phi\in S_{\emptyset,n}$. Hence

$$(2\pi)^{\frac{d}{2}}R_x\in G(\cdot-x)-\sum_{i=1}^{M}l_i(x)\,G(\cdot-a_i)+P_\theta,$$

so that for each x, $q_x\in P_\theta$ and

$$(2\pi)^{\frac{d}{2}}R_x=G(\cdot-x)-\sum_{i=1}^{M}l_i(x)\,G(\cdot-a_i)+q_x. \tag{4.33}$$

We now calculate q_x by applying the operator $\mathcal{P}$ to 4.33. This involves using the result $\mathcal{P}R_x=\sum\limits_{j=1}^{M}l_j(x)\,l_j$ and assuming that $q_x\in P_\theta$, so that $\mathcal{P}q_x=q_x$. In this way we obtain

$$q_x=\sum_{i=1}^{M}l_i(x)\,\mathcal{P}G(\cdot-a_i)-\mathcal{P}G(\cdot-x)+(2\pi)^{\frac{d}{2}}\sum_{j=1}^{M}l_j(x)\,l_j, \tag{4.34}$$

so that

$$\begin{aligned}(2\pi)^{\frac{d}{2}}R_x(y)\;=\;&G(y-x)-\sum_{i=1}^{M}l_i(x)\,G(y-a_i)-\sum_{j=1}^{M}G(a_j-x)\,l_j(y)+\\&+\sum_{i,j=1}^{M}l_i(x)\,G(a_j-a_i)\,l_j(y)+(2\pi)^{\frac{d}{2}}\sum_{j=1}^{M}l_j(x)\,l_j(y).\end{aligned} \tag{4.35}$$

This equation will be used to define $R_x(y)$ and noting 4.30, $D_x^\gamma R_x$ will be our candidate for Riesz representer for the evaluation functional $D^\gamma u(x)$ when $|\gamma|\le\lfloor\kappa\rfloor$.

Theorem 117 *Suppose G is a basis function of order θ generated by a weight function with properties W2 and W3. Suppose the operator $\mathcal{P}$ is defined with the same minimal unisolvent set used to define R_x in 4.35. Then*

$$R_x(y)=(2\pi)^{-\frac{d}{2}}\,\mathcal{Q}_y\mathcal{Q}_xG(y-x)+\sum_{j=1}^{M}l_j(x)\,l_j(y), \tag{4.36}$$

where $\mathcal{Q}=I-\mathcal{P}$.

Proof. The basis function is continuous by Theorems 49 or 50. Expanding $\mathcal{Q}_x$ and then $\mathcal{Q}_y$ yields equation 4.35 for R_x. ■

Theorem 118 $\mathcal{Q}_y\mathcal{Q}_x(y+x)^\alpha=0$ *when* $|\alpha|<2\theta$.

Proof. $\mathcal{Q}_y \mathcal{Q}_x (y+x)^\alpha = \mathcal{Q}_y \mathcal{Q}_x \sum\limits_{\beta \le \alpha} \binom{\alpha}{\beta} y^\beta (-x)^{\alpha-\beta} = \sum\limits_{\beta \le \alpha} \binom{\alpha}{\beta} \left(\mathcal{Q}_y y^\beta\right) \left(\mathcal{Q}_x x^{\alpha-\beta}\right)$, and since $|\alpha| = |\beta| + |\alpha - \beta| < 2\theta$ it follows that $|\beta| < \theta$ or $|\alpha-\beta| < \theta$ and hence that $\left(\mathcal{Q}_y y^\beta\right)\left(\mathcal{Q}_x x^{\alpha-\beta}\right) = 0$ for all $\beta \le \alpha$. ■

Theorem 119 *For each x, as a function:*

1. *R_x is independent of the basis function used to define it.*
2. *R_x is independent of the order of A.*

Proof. Part 1 From Definition 44 of a basis function, any two basis functions differ by a polynomial of order of at most 2θ. Our result now follows easily from Theorem 117 and Theorem 118.

Part 2 To prove this part we use equation 4.36 for $R_x(y)$. From Theorem 102 the function $\mathcal{Q}$ is independent of the order of the unisolvent set which generates it. Thus $(2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x G(y-x)$ is also independent of the order of the unisolvent set which generates it. Further, $\sum\limits_{j=1}^{M} l_j(x) l_j(y)$ is independent of the order of the cardinal basis functions. Hence R_x is independent of the order of A. ■

Theorem 120 *Suppose the weight function w has properties W2 and W3 for order θ and parameter κ. Suppose the distribution $R_x \in S'$ is defined by equation 4.35. Then for $|\gamma| \le \lfloor \kappa \rfloor$ we have $D_x^\gamma R_x \in X_w^\theta$ and*

$$\left|D_x^\gamma R_x\right|_{w,\theta}^2 = (2\pi)^{-d} \int \frac{\left|D_x^\gamma \mathcal{Q}_x \left(e^{-i(x,\cdot)}\right)\right|^2}{w\,|\cdot|^{2\theta}}, \quad x \in \mathbb{R}^d, \tag{4.37}$$

where 4.14 provides an upper bound for the right side of 4.37.

Proof. We prove that $D_x^\gamma R_x \in X_w^\theta$ by using definition 4.22 i.e. we show that $D_x^\gamma R_x \in S'$, $\widehat{D_x^\gamma R_x} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$, $\int w\,|\cdot|^{2\theta} \left|\left(D_x^\gamma R_x\right)_F\right|^2 < \infty$ and that $|\alpha| = \theta$ implies $\xi^\alpha \widehat{D_x^\gamma R_x} = \xi^\alpha \left(D_x^\gamma R_x\right)_F$ on S. If these criteria are satisfied then $\left|D_x^\gamma R_x\right|_{w,0}^2 = \int w\,|\cdot|^{2\theta} \left|\left(D_x^\gamma R_x\right)_F\right|^2$.

From 4.33, for each $x \in \mathbb{R}^d$

$$\begin{aligned}
(2\pi)^{d/2}\, \widehat{D_x^\gamma R_x} &= e^{-i(x,\cdot)} \widehat{D^\gamma G} - \sum_{i=1}^{M} D^\gamma l_i(x)\, e^{-i\langle a_i,\cdot\rangle} \widehat{G} + \widehat{q_x} \\
&= e^{-i(x,\cdot)} (-i\xi)^\gamma\, \widehat{G} - \sum_{i=1}^{M} D^\gamma l_i(x)\, e^{-i\langle a_i,\cdot\rangle} \widehat{G} + \widehat{q_x} \\
&= \left(e^{-i(x,\cdot)} (-i\xi)^\gamma - \sum_{i=1}^{M} D^\gamma l_i(x)\, e^{-i\langle a_i,\cdot\rangle}\right) \widehat{G} + \widehat{q_x} \\
&= D_x^\gamma \left(e^{-i(x,\cdot)} - \sum_{i=1}^{M} l_i(x)\, e^{-i\langle a_i,\cdot\rangle}\right) \widehat{G} + \widehat{q_x} \\
&= \left(D_x^\gamma \mathcal{Q}_x \left(e^{-i(x,\cdot)}\right)\right) \widehat{G} + \widehat{q_x},
\end{aligned} \tag{4.38}$$

where $q_x \in P_\theta$. Since W2 implies $1/w \in L_{loc}^1$, Definition 44 of a basis distribution implies $\widehat{G} = \frac{1}{w|\cdot|^{2\theta}} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ and so 4.38 implies that for each x

$$\left(D_x^\gamma R_x\right)_F = (2\pi)^{-\frac{d}{2}} \frac{D_x^\gamma \mathcal{Q}_x \left(e^{-i(x,\cdot)}\right)}{w\,|\cdot|^{2\theta}} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right). \tag{4.39}$$

Again by 4.38 and Theorem 110

$$\int w\,|\cdot|^{2\theta} \left|\left(D_x^\gamma R_x\right)_F\right|^2 = (2\pi)^{-d} \int \frac{\left|D_x^\gamma \mathcal{Q}_x \left(e^{-i(x,\cdot)}\right)\right|^2}{w\,|\cdot|^{2\theta}} < \infty, \tag{4.40}$$

Now use ξ as an *action* variable. It must be shown that $\xi^\alpha \widehat{D_x^\gamma R_x} = \xi^\alpha \left(D_x^\gamma R_x\right)_F$ on S when $|\alpha| = \theta$, which is true if $\xi^\alpha \widehat{R_x} = \xi^\alpha \left(R_x\right)_F$ on S when $|\alpha| = \theta$. But from 4.38 it follows that $\xi^\alpha \widehat{R_x} =$

$\xi^\alpha \mathcal{Q}_x \left(e^{-i(x,\xi)}\right) \widehat{G} + \xi^\alpha \widehat{q_x} = \xi^\alpha \mathcal{Q}_x \left(e^{-i(x,\xi)}\right) \widehat{G}$ since $q_x \in P_\theta$. From part 4 of Theorem 106, if $|\alpha| = \theta$ then for each x,

$(i\xi)^\alpha \mathcal{Q}_x \left(e^{i(x,\xi)}\right) \in C^\infty_{\emptyset,2\theta} \cap C^\infty_{BP}$ and so $\psi \in S$ implies $(i\xi)^\alpha \mathcal{Q}_x \left(e^{i(x,\xi)}\right) \psi \in S_{\emptyset,2\theta}$ by Theorem 15. Thus by the definition of a basis distribution

$$\left[\xi^\alpha \widehat{R_x}, \psi\right] = \left[\xi^\alpha \mathcal{Q}_x \left(e^{-i(x,\xi)}\right) \widehat{G}, \psi\right] = \left[\widehat{G}, \xi^\alpha \mathcal{Q}_x \left(e^{-i(x,\xi)}\right) \psi\right] = \int \frac{\xi^\alpha \mathcal{Q}_x \left(e^{-i(x,\xi)}\right)}{w(\xi) |\xi|^{2\theta}} \psi(\xi)\, d\xi$$
$$= \int \xi^\alpha (R_x)_F \psi,$$

and hence $\xi^\alpha \widehat{R_x} = \xi^\alpha (R_x)_F$ on S. We can now conclude that $D^\gamma_x R_x \in X^\theta_w$, that 4.40 is true and so 4.37 is proven. ■

Corollary 121 *Suppose the weight function w has properties W2 and W3 for order θ and smoothness parameter κ. Then for $|\delta|, |\gamma| \le \lfloor\kappa\rfloor$ we have*

$$\left\langle D^\delta_y R_y, D^\gamma_x R_x \right\rangle_{w,\theta} = (2\pi)^{-d} \int \frac{D^\delta_y \mathcal{Q}_y \left(e^{-i(y,\cdot)}\right) \overline{D^\gamma_x \mathcal{Q}_x \left(e^{-i(x,\cdot)}\right)}}{w |\cdot|^{2\theta}}, \quad x, y \in \mathbb{R}^d,$$

Proof. Write $\left\langle D^\delta_y R_y, D^\gamma_x R_x \right\rangle_{w,\theta} = \int w |\cdot|^{2\theta} \left(D^\delta_y R_y\right)_F \overline{(D^\gamma_x R_x)_F}$ and then use 4.39. ■

Now we prove the important result that when $|\gamma| \le \kappa$, $D^\gamma_x R_x$ is the Riesz representer for the evaluation functional $u \to D^\gamma u(x)$.

Theorem 122 *If $u \in X^\theta_w$ and $|\gamma| \le \kappa$ then*

$$D^\gamma u(x) = (u, D^\gamma_x R_x)_{w,\theta}, \quad x \in \mathbb{R}^d,$$

where R_x and the Light norm 4.29 are defined using the same minimal θ-unisolvent set $A = \{a_k\}_{k=1}^M$.

Proof. Suppose R_x and the Light norm are defined using the minimal unisolvent set $A = \{a_k\}_{k=1}^M$ and that the corresponding cardinal basis $\{l_k\}_{k=1}^M$ of P_θ. From the definition of the Light norm

$$(u, D^\gamma_x R_x)_{w,\theta} = \int w |\cdot|^{2\theta} u_F \overline{(D^\gamma_x R_x)_F} + \sum_{k=1}^M u(a_k) \overline{(D^\gamma_x R_x)(a_k)},$$

where the function $(D^\gamma_x R_x)_F$ is given by 4.39. Hence

$$\int w |\cdot|^{2\theta} u_F \overline{(D^\gamma_x R_x)_F} = (2\pi)^{-\frac{d}{2}} \int w |\cdot|^{2\theta} u_F \frac{D^\gamma_x \mathcal{Q}_x \left(e^{i(x,\cdot)}\right)}{w |\cdot|^{2\theta}} = (2\pi)^{-\frac{d}{2}} \int u_F D^\gamma_x \mathcal{Q}_x \left(e^{i(x,\cdot)}\right),$$

and by 4.32, $D^\gamma l_k(x) = (D^\gamma_x R_x)(a_k)$ so that

$$\sum_{k=1}^M u(a_k) \overline{(D^\gamma_x R_x)(a_k)} = \sum_{k=1}^M u(a_k) D^\gamma l_k(x) = (D^\gamma \mathcal{P} u)(x).$$

Thus

$$(u, D^\gamma_x R_x)_{w,\theta} = (2\pi)^{-\frac{d}{2}} \int u_F D^\gamma_x \mathcal{Q}_x \left(e^{i(x,\cdot)}\right) + (D^\gamma \mathcal{P} u)(x),$$

and comparison with equation 4.23 of Theorem 113 for $D^\gamma u$ implies $D^\gamma u(x) = (u, D^\gamma_x R_x)_{w,\theta}$ as required. ■

Showing that X^θ_w is a reproducing kernel Hilbert space now becomes very simple.

Corollary 123 *Suppose the weight function w has properties W2 and W3 for order θ and parameter κ. Then X^θ_w is a reproducing kernel Hilbert space when endowed with the Light norm.*

Proof. By Theorem 120 $R_x \in X^\theta_w$. Hence by Theorem 122 with $\gamma = 0$ we have $|u(x)| \le \left|(u, R_x)_{w,\theta}\right| \le \|R_x\|_{w,\theta} \|u\|_{w,\theta}$ and $\|R_x\|_{w,\theta}$ is a finite constant for each x. Hence X^θ_w is a reproducing kernel Hilbert space e.g. Theorem III.9.1, Yosida [23]. ■

Theorem 124 ***Properties of*** $D_x^\gamma R_x$
Suppose the weight function w satisfies properties W2 and W3 for order θ and κ. Suppose R_x is defined using the minimal unisolvent set $A = \{a_j\}_{j=1}^M$. Then for all $x, y \in \mathbb{R}^d$ and $|\gamma| \le \lfloor \kappa \rfloor$:

1. $R_x(y) = \overline{R_y(x)}$.
2. *For* $j = 1, \ldots, M$, $(D_x^\gamma R_x)(a_j) = D^\gamma l_j(x)$, *and each cardinal basis polynomial* l_j *has real coefficients.*
3. $D^\gamma \mathcal{Q}u(x) = \langle u, D_x^\gamma R_x \rangle_{w,\theta}$, $u \in X_w^\theta$.
4. $\mathcal{P}D_x^\gamma R_x = \sum\limits_{j=1}^M D^\gamma l_j(x)\, l_j = D_x^\gamma \mathcal{P}R_x$ *and* $\mathcal{Q}D_x^\gamma R_x = D_x^\gamma \mathcal{Q}R_x$.
5. *The Riesz representer is unique.*

Proof. Parts 1 and **2** By Theorem 122 $D_x^\gamma R_x$ is the Riesz representer of the evaluation functional $u \to D^\gamma u(x)$ and so equations 4.31 and 4.32 are valid. By Definition 100 the cardinal basis polynomials have real coefficients.

Part 3 If $u \in X_w^\theta$ then by part 1 of Theorem 116 $D^\gamma \mathcal{Q}u(x) = (\mathcal{Q}u, D_x^\gamma R_x)_{w,\theta} = \langle u, D_x^\gamma R_x \rangle_{w,\theta}$.

Part 4 Using part 2

$$\mathcal{P}D_x^\gamma R_x = \sum_{j=1}^M (D_x^\gamma R_x)(a_j)\, l_j = \sum_{j=1}^M (D^\gamma l_j)(x)\, l_j = D_x^\gamma \sum_{j=1}^M l_j(x)\, l_j = D_x^\gamma \mathcal{P}R_x,$$

so that

$$\mathcal{Q}D_x^\gamma R_x = D_x^\gamma R_x - \mathcal{P}D_x^\gamma R_x = D_x^\gamma R_x - D_x^\gamma \mathcal{P}R_x = D_x^\gamma \mathcal{Q}R_x.$$

Part 5 Suppose two representers exist, say $D_x^\gamma R_x$ and $S_x^{(\gamma)}$. Then for all $u \in X_w^\theta$ and $x \in \mathbb{R}^d$, $(u, D_x^\gamma R_x)_{w,\theta} = \left(u, S_x^{(\gamma)}\right)_{w,\theta} = D^\gamma u(x)$ and hence
$\left(u, S_x^{(\gamma)} - D_x^\gamma R_x\right)_{w,\theta} = 0$ which implies $S_x^{(\gamma)} = D_x^\gamma R_x$. ∎

4.5.4 The semi-Riesz representer $r_x = \mathcal{Q}R_x$

Here we will introduce what I will call the *semi-Riesz representer* $r_x = \mathcal{Q}R_x$ for the evaluation functional $u \to u(x)$. This terminology is based on the equation $\mathcal{Q}u(x) = \langle u, r_x \rangle_{w,\theta}$, $u \in X_w^\theta$, derived in part 7 of the next theorem. A bound for the the function $r_x(x)$ will be obtained in Subsection 4.10.2 which will then be used to estimate the pointwise rate of convergence of the interpolant.

Theorem 125 *Suppose the function $r_x = \mathcal{Q}R_x$ is defined using the minimal unisolvent set $\{a_k\}_{k=1}^M$ and corresponding cardinal basis $\{l_k\}_{k=1}^M$ of P_θ. Then r_x has the properties:*

1. $r_x(y) = R_x(y) - \sum\limits_{j=1}^M l_j(x) l_j(y)$.
2. $r_x(y) = \overline{r_y(x)}$.
3. $r_x(a_i) = r_{a_i}(x) = 0$ *for all* i.
4. $\mathcal{P}r_x = 0$ *and* $\mathcal{Q}r_x = r_x$.
5. $r_x(y) = \langle r_x, r_y \rangle_{w,\theta}$.
6. $r_x(y) = (2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x G(y - x)$.
7. $\mathcal{Q}u(x) = \langle u, r_x \rangle_{w,\theta}$ *when* $u \in X_w^\theta$.

Proof. Part 1 Apply $\mathcal{Q}_y$ to equation 4.36. **Part 2** Follows from part 1 since $R_x(y) = \overline{R_y(x)}$ and the l_k are real valued. **Part 3** Part 2 of Theorem 124 with $\gamma = 0$ implies $R_x(a_i) = l_i(x)$. Parts 1 and 2 now give this result. **Part 4** Use part 3. **Part 5** Part 3 of Theorem 124 with $u = r_x$ and $\gamma = 0$ gives $\mathcal{Q}r_x(y) = \langle r_x, R_y \rangle_{w,\theta}$. Since $\mathcal{Q}r_x = r_x$, Theorem 116 implies $r_x(y) = \mathcal{Q}r_x(y) = \langle r_x, R_y \rangle_{w,\theta} = \langle r_x, \mathcal{Q}R_y \rangle_{w,\theta} = \langle r_x, r_y \rangle_{w,\theta}$.

Part 6 Apply $\mathcal{Q}_y$ to formula 4.36. **Part 7** follows from part 3 of Theorem 124 with $\gamma = 0$, and then use part 4. ∎

Remark 126 *The projection properties of the operators $\mathcal{P}$ and $\mathcal{Q}$ can be used to show that $X_w^\theta = \mathcal{Q}\left(X_w^\theta\right) \oplus P_\theta$ and that $\mathcal{Q}\left(X_w^\theta\right)$ is a Hilbert space when endowed with the semi-inner product $\langle \cdot, \cdot \rangle_{w,\theta}$. In fact, on $\mathcal{Q}\left(X_w^\theta\right)$, r_x is the Riesz representer of the functional $f \to f(x)$ and $\mathcal{Q}\left(X_w^\theta\right)$ is a reproducing kernel Hilbert space.*

4.6 The basis function and reproducing kernel matrices

In this section we will introduce the *basis function matrix* and the *reproducing kernel matrix*. Together with the unisolvency matrices these matrices will be used to construct matrix equations for the basis function interpolants of this document.

4.6.1 The basis function matrix $G_{X,X}$

The *basis function matrix* and the unisolvency matrix will be used to construct the block matrix equation 4.66 for the variational interpolation in Subsection 4.9.4.

Definition 127 ***The basis function matrix***

Let $X = \left\{x^{(n)}\right\}_{n=1}^N$ be N distinct points in $\mathbb{R}^d$ and suppose G is a (continuous) basis function. Then the $N \times N$ basis function matrix $G_{X,X}$ is defined by

$$G_{X,X} = \left(G\left(x^{(i)} - x^{(j)}\right)\right).$$

4.6.2 The reproducing kernel matrix $R_{X,X}$

The *reproducing kernel matrix* is derived from the Riesz representer of the evaluation functional $u \to u(x)$ discussed in Section 4.5.3. I call it the reproducing kernel matrix because the existence of this special evaluation functional means that the space X_w^θ is a reproducing kernel Hilbert space when endowed with the Light norm. This matrix is closely linked to the basis function matrix of the previous subsection will be used to construct a matrix equation for the variational interpolation problems of Section 4.9. In the next chapter we will study the relationships between the reproducing kernel matrix and the basis function matrix and derive a matrix equation for the Exact smoother using the reproducing kernel matrix.

Definition 128 ***The reproducing kernel matrix*** $R_{X,X}$

Suppose a basis function has order θ and suppose that $R_x(y)$ is the Riesz representer of the evaluation functional $u \to u(x)$, $u \in X_w^\theta$ defined by 4.35. Also suppose that $X = \left\{x^{(n)}\right\}_{n=1}^N$ is a set of distinct unisolvent data points in $\mathbb{R}^d$. Then define the $N \times N$ matrix $R_{X,X}$ by

$$R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right).$$

Theorem 129 *Suppose the conditions imposed in the definition of the reproducing kernel matrix $R_{X,X}$ hold. Then the following are true:*

1. *$R_{X,X} = \left(\left(R_{x^{(j)}}, R_{x^{(i)}}\right)_{w,\theta}\right)$ so it is a Gram matrix i.e. has inner product elements $\left(u_i, u_j\right)_{w,\theta}$.*

2. *The matrix $R_{X,X}$ is Hermitian, regular and positive definite.*

3. *The functions $\left\{R_{x^{(i)}}\right\}_{i=1}^N$ are linearly independent.*

Proof. Part 1 By definition of R_x, $R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right) = \left(R_{x^{(j)}}, R_{x^{(i)}}\right)_{w,\theta}$.

Part 2 These are elementary properties of a Gram matrix.

Part 3 This is a direct consequence of the regularity of $R_{X,X}$. ∎

4.7 The basis function spaces $\dot{W}_{G,X}$ and $W_{G,X}$

The importance of the finite dimensional spaces $\dot{W}_{G,X}$ and $W_{G,X}$ is that they will contain the basis function interpolants of this document and the basis function smoothers of the succeeding documents. Here X is the independent data and will be assumed to be unisolvent. The finite dimensionality of these spaces means that matrix equations can be constructed for the interpolant and the smoothers.

Definition 130 ***The basis function spaces*** $W_{G,X}$ ***and*** $\dot{W}_{G,X}$

Suppose the weight function w has properties W2 and W3 for order θ and smoothness parameter κ. Then the basis distributions of order θ are continuous functions and let G be a basis function. Let $X = \{x^{(i)}\}_{i=1}^{N}$ be a θ-unisolvent set of distinct points in $\mathbb{R}^d$ and set $M = \dim P_\theta$. Next choose a (real) basis $\{p_j\}_{j=1}^{M}$ of P_θ and calculate the unisolvency matrix $P_X = \left(p_j\left(x^{(i)}\right)\right)$. We can now define

$$\dot{W}_{G,X} = \left\{ \sum_{i=1}^{N} v_i G\left(x - x^{(i)}\right) : v_i \in \mathbb{C} \text{ and } P_X^T v = 0 \right\}, \tag{4.41}$$

$$W_{G,X} = \dot{W}_{G,X} + P_\theta.$$

Clearly $\dot{W}_{G,X}$ and $W_{G,X}$ are vector spaces. When convenient, functions in spaces of the form $\dot{W}_{G,X}$ will be written, $f_v(x) = \sum_{i=1}^{N} v_i G\left(x - x^{(i)}\right)$.

In words, $\dot{W}_{G,X}$ is a ***subspace*** *of the complex span of the $X-$translated basis functions and $W_{G,X}$ is a* ***subspace*** *of the complex span of the $X-$translated basis functions plus the polynomials of at most basis function order. Sometimes we will say data-translated basis functions.*

The next theorem shows that the set $\dot{W}_{G,X}$ is independent of the order of the points in X and of the basis of P_θ used to calculate P_X.

Theorem 131 *Let $\{p_i\}$ and $\{q_i\}$ be two bases for the polynomials P_θ, and set $P_X = \left(p_j\left(x^{(i)}\right)\right)$ and $Q_X = \left(q_j\left(x^{(i)}\right)\right)$. Then*

$$\left\{ \sum_{i=1}^{N} v_i G\left(x - x^{(i)}\right) : P_X^T v = 0 \right\} = \left\{ \sum_{i=1}^{N} \mu_i G\left(x - x^{(i)}\right) : Q_X^T \mu = 0 \right\}, \tag{4.42}$$

and $\dot{W}_{G,X}$ is independent of the basis for P_θ used to define P_X.

Proof. From part 3 of Theorem 104 there exists a regular matrix R such that $P_X = Q_X R^T$. Hence $P_X^T = R Q_X^T$ and then equation 4.42 easily follows. ■

Now recall Definition 100 regarding permutation theory.

Theorem 132 *As a set $\dot{W}_{G,X}$ is independent of the ordering of X and independent of the basis function G used to define it.*

Proof. Using the notation $G_{x,X} = \left(G\left(x - x^{(j)}\right)\right)$ we have

$$\dot{W}_{G,X} = \left\{ \sum_{i=1}^{N} v_i G\left(x - x^{(i)}\right) : P_X^T v = 0 \right\} = \left\{ G_{x,X} v : P_X^T v = 0 \right\},$$

where $P_X^T = \left(p_i\left(x^{(j)}\right)\right) = \begin{pmatrix} p_1\left(x^{(1)}\right) & p_1\left(x^{(2)}\right) & \cdots & p_1\left(x^{(N)}\right) \\ p_2\left(x^{(1)}\right) & p_2\left(x^{(2)}\right) & \cdots & p_2\left(x^{(N)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ p_M\left(x^{(1)}\right) & p_M\left(x^{(2)}\right) & \cdots & p_M\left(x^{(N)}\right) \end{pmatrix}$.

Suppose re-ordering X uses the permutation π. Then re-ordering a row vector involves right multiplication by the permutation matrix Π^T so that $G_{x,\pi(X)} = G_{x,X}\Pi^T$, $P_{\pi(X)}^T = P_X^T \Pi^T$ and hence

$$\dot{W}_{G,\pi(X)} = \left\{ G_{x,\pi(X)} : P_{\pi(X)}^T v = 0 \right\} = \left\{ G_{x,\pi(X)} \Pi^T v : P_X^T \Pi^T v = 0 \right\} = \dot{W}_{G,X}.$$

Set-wise independence of the basis function will hold if $p \in P_\theta$ implies $\sum_{i=1}^{N} v_i p\left(x - x^{(i)}\right) = 0$ and clearly we need only show this for the basis functions $\{x^\alpha\}_{|\alpha|<\theta}$. Since $\{x^\alpha\}_{|\alpha|<\theta}$ is a basis, by Theorem 131 $P_X^T v = 0$ iff $\sum_{i=1}^{N} v_i \left(x^{(i)}\right)^\alpha = 0$ for $|\alpha| < \theta$. Thus

$$\sum_{i=1}^{N} v_i \left(x - x^{(i)}\right)^\alpha = \sum_{i=1}^{N} v_i \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^\beta \left(x^{(i)}\right)^{\alpha-\beta} = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^\beta \sum_{i=1}^{N} v_i \left(x^{(i)}\right)^{\alpha-\beta} = 0,$$

confirming set-wise independence. ■

Lemma 133 *Suppose $X = \{x_k\}_{k=1}^N$ is θ-unisolvent and P_X is a unisolvency matrix of order θ (see Definition 103). Further, suppose $v = (v_k) \in \mathbb{R}^N$ satisfies $P_X^T v = 0$. Now define the function $a_v : \mathbb{R}^d \to \mathbb{C}$ by*

$$a_v(z) = \sum_{k=1}^{N} v_k e^{-i(x_k, z)}, \quad z \in \mathbb{R}^d. \tag{4.43}$$

Then the function a_v has the following properties:

1. *$a_v \in C_{\emptyset,\theta}^\infty \cap C_B^\infty$.*

2. *If $\phi \in S$ then $|a_v|^2 \phi \in S$ and*

$$\widehat{|a_v|^2 \phi} = \sum_{j,k=1}^{N} v_j v_k \widehat{\phi}\left(\cdot - (x_k - x_j)\right).$$

3. *$a_v(z) = 0$ a.e. implies $v_k = 0$ for all k.*

Proof. Part 1 Since

$$D^\beta a_v(z) = D^\beta \sum_{k=1}^{N} v_k e^{-i(x_k, z)} = \sum_{k=1}^{N} v_k (-i x_k)^\beta e^{-i(x_k, z)},$$

it is clear that all derivatives are bounded and hence $a_v \in C_B^\infty$. If $|\beta| < \theta$ then

$$D^\beta a_v(0) = \sum_{k=1}^{N} v_k (-i x_k)^\beta = (-i)^{|\beta|} \sum_{k=1}^{N} v_k x_k^\beta = 0,$$

since $P_X^T v = 0$.

Part 2

$$\begin{aligned} |a_v(x)|^2 \phi(x) &= \left| \sum_{j=1}^{N} v_j e^{-i(x_j, x)} \right|^2 \phi(x) = \left(\sum_{j=1}^{N} v_j e^{-i(x_j, x)} \right) \left(\overline{\sum_{k=1}^{N} v_k e^{-i(x_k, x)}} \right) \phi(x) \\ &= \sum_{j,k=1}^{N} v_j v_k e^{-i(x_j - x_k, x)} \phi(x). \end{aligned}$$

Therefore

$$\begin{aligned} \widehat{|a_v|^2 \phi}(\eta) = \sum_{j,k=1}^{N} v_j v_k \left[e^{-i(x_j - x_k, x)} \phi(x) \right]^\wedge(\eta) &= \sum_{j,k=1}^{N} v_j v_k \widehat{\phi}(x_j - x_k + \eta) \\ &= \sum_{j,k=1}^{N} v_j v_k \widehat{\phi}\left(\eta - (x_k - x_j)\right). \end{aligned}$$

Part 3 Let $\Delta = \{x_k - x_j : 1 \le j,k \le N,\ k \ne j\}$. Then $0 \notin \Delta$ and so $\delta = \operatorname{dist}(0;\Delta) > 0$. Next let ω_ε be the standard 'cap-shaped' distribution test function

$$\omega_\varepsilon(\xi) = \begin{cases} C_\varepsilon \exp\left(-\frac{\varepsilon^2}{\varepsilon^2 - |\xi|^2}\right), & |\xi| \le \varepsilon, \\ 0, & |\xi| > \varepsilon, \end{cases}$$

with C_ε chosen so that $\omega_\varepsilon(0) = 1$. Now ω_ε has support $\overline{B(0;\varepsilon)}$ and if we set $\widehat{\phi}(\xi) = \omega_\delta(\xi)$ it follows that $\widehat{\phi}(0) = 1$ and $\widehat{\phi}(\xi) = 0$ when $\xi \in \Delta$. Thus $a_v = 0$ a.e. implies $\widehat{|a_v|^2 \phi} = 0$ a.e. and so, $0 = \sum\limits_{j,k=1}^{N} v_j v_k \widehat{\phi}(\xi - (x_k - x_j)) = \sum\limits_{k=1}^{N} v_k^2$ and therefore $v_k = 0$ for all k. ■

Next we obtain expressions for the semi-inner product and seminorm of functions in $\dot{W}_{G,X}$.

Theorem 134 *Suppose w is a weight function with properties W2.1 and W3 for some order $\theta \ge 1$. Suppose $v = (v_i)_{i=1}^N$ is a complex vector satisfying $P_X^T v = 0$ where P_X is the unisolvency matrix. Then we have the following results:*

1. *If the functions f_v and a_v are defined by*

$$f_v(x) = \sum_{k=1}^{N} v_k G\left(x - x^{(k)}\right) \in \dot{W}_{G,X}, \quad a_v(\xi) = \sum_{j=1}^{N} v_j e^{-i\left(x^{(j)},\xi\right)}, \tag{4.44}$$

then $f_v \in X_w^\theta$ and

$$|f_v|_{w,\theta}^2 = \int \frac{|a_v|^2}{w\,|\cdot|^{2\theta}} = (2\pi)^{\frac{d}{2}} \sum_{j,k=1}^{N} v_j \overline{v_k} G\left(x^{(j)} - x^{(k)}\right) = (2\pi)^{\frac{d}{2}} v^T G_{X,X} \overline{v}. \tag{4.45}$$

2. *Each $f \in \dot{W}_{G,X}$ has a unique representation 4.41 and* $\dim \dot{W}_{G,X} = N - M$.
3. *$G_{X,X}$ is conditionally positive definite on* null P_X^T *i.e. $P_X^T v = 0$ and $v \ne 0$ implies $v^T G_{X,X} \overline{v} > 0$.*
4. *We have the direct sum $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$ and* $\dim W_{G,X} = N$.
5. *$X_w^\theta = W_{G,X} \oplus W_{G,X}^\perp$ where*

$$W_{G,X}^\perp = \left\{ u \in X_w^\theta : u\left(x^{(k)}\right) = 0 \textit{ for all } x^{(k)} \in X \right\}.$$

Proof. Part 1 The first step is to show that $f_v \in X_w^\theta$. Here we use the definition of X_w^θ given in equation 4.22 i.e. $f_v \in S'$, $\widehat{f_v} \in L_{loc}^1(\mathbb{R}^d \backslash 0)$, $\int w\,|\cdot|^{2\theta} |(f_v)_F|^2 < \infty$ and $|\alpha| = \theta$ implies $\xi^\alpha \widehat{f_v} = \xi^\alpha (f_v)_F$ on S. Clearly since $G \in S'$ we have $f_v \in S'$ and

$$\widehat{f_\nu} = \sum_{j=1}^{N} v_j F_x\left[G\left(x - x^{(j)}\right)\right] = \sum_{j=1}^{N} v_j e^{-i\left(x^{(j)},\xi\right)} \widehat{G} = \left(\sum_{j=1}^{N} v_j e^{-i\left(x^{(j)},\xi\right)}\right) \widehat{G} = a_v(\xi)\,\widehat{G},$$

where a_v is given by 4.44.
Thus $\widehat{f_\nu} = \frac{a_v}{w|\cdot|^{2\theta}}$ on $\mathbb{R}^d \backslash 0$ and since property W2.1 is $1/w \in L_{loc}^1$ it follows that $\widehat{f_\nu} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ and

$$(f_\nu)_F = \frac{a_v}{w\,|\cdot|^{2\theta}}. \tag{4.46}$$

By Theorem 133, $a_v \in C_{\emptyset,\theta}^\infty \cap C_B^\infty$ and so by Theorem 15 we have $|a_v|^2 \in C_{\emptyset,2\theta}^\infty \cap C_B^\infty$. Further

$$|f_\nu|_{w,\theta}^2 = \int w\,|\cdot|^{2\theta} |(f_\nu)_F|^2 = \int w\,|\cdot|^{2\theta} \left|\frac{a_v}{w\,|\cdot|^{2\theta}}\right|^2 = \int \frac{|a_v|^2}{w\,|\cdot|^{2\theta}}. \tag{4.47}$$

From Theorem 10, W3.1 implies property W3.2 and by definition property W3.2 implies $\int_{|\cdot|\geq r_3} \frac{1}{w|\cdot|^{2\theta}} < \infty$ for some $r_3 \geq 0$. Choose $\phi \in C_0^\infty$ such that $0 \leq \phi \leq 1$ and $\phi(x) = 1$ in a neighborhood of 0, and define ϕ_ε by: $\phi_\varepsilon(\xi) = \phi(\varepsilon\xi)$ for $\varepsilon > 0$. Then

$$\int \frac{|a_v|^2}{w|\cdot|^{2\theta}} = \int \frac{\phi_\varepsilon |a_v|^2}{w|\cdot|^{2\theta}} + \int \frac{(1-\phi_\varepsilon)|a_v|^2}{w|\cdot|^{2\theta}}, \tag{4.48}$$

and since $\phi_\varepsilon |a_v|^2 \in S_{\emptyset,2\theta}$ we have by the basis function Definition 44 that $\int \frac{\phi_\varepsilon|a_v|^2}{w|\cdot|^{2\theta}} = \left[\widehat{G}, \phi_\varepsilon |a_v|^2\right] < \infty$. Further, for sufficiently small ε the support of $(1-\phi_\varepsilon)|a_v|^2$ lies outside the sphere of radius r_3 so property W3.2 now ensures that the integral 4.48 exists. To finish the proof that $f_v \in X_w^\theta$ we must show that $\xi^\alpha (f_\nu)_F = \xi^\alpha \widehat{f_\nu}$ on S i.e. $\xi^\alpha \frac{a_v}{w|\cdot|^{2\theta}} = \xi^\alpha a_v \widehat{G}$ on S.

If $|\alpha| = \theta$ then by Theorem 15, $\xi^\alpha \in C_{\emptyset,\theta}^\infty$, $\xi^\alpha a_v \in C_{\emptyset,2\theta}^\infty \cap C_{BP}^\infty$ and $\xi^\alpha a_v \psi \in S_{\emptyset,2\theta}$ when $\psi \in S$ so that

$$\left[\xi^\alpha \frac{a_v}{w|\cdot|^{2\theta}}, \psi\right] = \left[\frac{1}{w|\cdot|^{2\theta}}, \xi^\alpha a_v \psi\right] = \left[\widehat{G}, \xi^\alpha a_v \psi\right] = \left[\xi^\alpha a_v \widehat{G}, \psi\right],$$

as required. Thus $f_v \in X_w^\theta$ and 4.47 proves the first equation of 4.45.

We have already shown that $|a_v|^2 \in C_{\emptyset,2\theta}^\infty \cap C_B^\infty$ so $|a_v|^2 \phi_\varepsilon \in S_{\emptyset,2\theta}$ and by definition of a basis distribution

$$\begin{aligned}
\int \frac{|a_v|^2}{w|\cdot|^{2\theta}}\phi_\varepsilon = \left[\frac{1}{w|\cdot|^{2\theta}}, |a_v|^2 \phi_\varepsilon\right] = \left[\widehat{G}, |a_v|^2 \phi_\varepsilon\right] &= \left[G, \widehat{|a_v|^2 \phi_\varepsilon}\right] \\
&= \sum_{j,k=1}^N v_j \overline{v_k} \left[G, \widehat{\phi_\varepsilon}\left(\cdot - \left(x^{(k)} - x^{(j)}\right)\right)\right] \\
&= \sum_{j,k=1}^N v_j \overline{v_k} \left[G, \overset{\vee}{\phi_\varepsilon}\left(\left(x^{(k)} - x^{(j)}\right) - \cdot\right)\right],
\end{aligned}$$

by part 2 of Theorem 133. By the convolution definition of part 2 Appendix 245,

$$\sum_{j,k=1}^N v_j \overline{v_k} \left[G, \overset{\vee}{\phi_\varepsilon}\left(\left(x^{(k)} - x^{(j)}\right) - \cdot\right)\right] = (2\pi)^{\frac{d}{2}} \sum_{j,k=1}^N v_j \overline{v_k} \left(G * \overset{\vee}{\phi_\varepsilon}\right)\left(x^{(k)} - x^{(j)}\right).$$

Now in the tempered distribution sense as $\varepsilon \to 0$, $\phi_\varepsilon \to 1$ and so $\overset{\vee}{\phi_\varepsilon} \to \overset{\vee}{1} = (2\pi)^{\frac{d}{2}}\delta$ and hence $G * \overset{\vee}{\phi_\varepsilon} \to G * (2\pi)^{\frac{d}{2}}\delta = G$ pointwise because by Theorem 49 and 50, if w has properties W2.1 and W3 then $G \in C^{(0)}$. Here $\overset{\vee}{\phi_\varepsilon}$ is called a mollifier e.g. Lemma 2.18 Adams [2]. Thus

$$\lim_{\varepsilon\to 0} \int \frac{|a_v|^2}{w|\cdot|^{2\theta}}\phi_\varepsilon = (2\pi)^{\frac{d}{2}} \sum_{j,k=1}^N v_j \overline{v_k} G\left(x^{(k)} - x^{(j)}\right),$$

and finally we note that $\lim\limits_{\varepsilon\to 0} \int \frac{|a_v|^2}{w|\cdot|^{2\theta}}\phi_\varepsilon = \int \frac{|a_v|^2}{w|\cdot|^{2\theta}}$.

Part 2 Suppose $\sum\limits_{k=1}^N v_k G\left(x - x^{(k)}\right) = 0$ for all x and $P_X^T v = 0$. Then

$$0 = \left|\sum_{k=1}^N v_k G\left(x - x^{(k)}\right)\right|_{w,\theta}^2 = \int \frac{|a_v|^2}{w|\cdot|^{2\theta}}.$$

Since $w > 0$ a.e. we must conclude that $a_v = 0$ a.e. and that by part 3 of Lemma 133, $v_k = 0$ for all k. Thus the operator $\Phi : \mathrm{null}\, P_X^T \to \dot{W}_{G,X}$ defined by $\Phi v = \sum_{k=1}^N v_k G\left(\cdot - x^{(k)}\right)$ is an isomorphism and so $\dim \dot{W}_{G,X} = \dim \mathrm{null}\, P_X^T = N - M$ by part 5 Theorem 104.

Part 3 From part 1, $P_X^T v = 0$ implies $v^T G_{X,X} \overline{v} \geq 0$. Suppose $v^T G_{X,X} \overline{v} = 0$ and $P_X^T v = 0$. By part 1 of this theorem

$$0 = v^T G_{X,X} \overline{v} = \left| \sum_{k=1}^{N} v_k G\left(x - x^{(k)}\right) \right|_{w,\theta},$$

so that $\sum\limits_{k=1}^{N} v_k G\left(x - x^{(k)}\right) = 0$. Part 2 of this theorem then implies $v = 0$.

Part 4 We must show that $\dot{W}_{G,X} \cap P_\theta = \{0\}$. Suppose $p \in \dot{W}_{G,X} \cap P_\theta$. Then $p = \sum\limits_{k=1}^{N} v_k G\left(x - x^{(k)}\right)$ and $P_X^T v = 0$ but by part 1, $0 = \|p\|_{w,\theta} = (2\pi)^{\frac{d}{2}} v^T G_{X,X} \overline{v}$. Part 2 now implies $v = 0$ and hence that $p = 0$.

Part 5 Since $\left\{R_{x^{(k)}} : x^{(k)} \in X\right\}$ is a basis for $W_{G,X}$

$$\begin{aligned} W_{G,X}^{\perp} = \left\{v \in X_w^\theta \mid (v,u)_{w,\theta} = 0 \ if \ u \in W_{G,X}\right\} &= \left\{v \in X_w^\theta \mid (v, R_{x^{(k)}})_{w,\theta} = 0 \ if \ x^{(k)} \in X\right\} \\ &= \left\{v \in X_w^\theta \mid v\left(x^{(k)}\right) = 0 \ if \ x^{(k)} \in X\right\}. \end{aligned}$$

■

A consequence of the last theorem is the following representation result for members of the set $W_{G,X}$:

Corollary 135 *Suppose the definitions and assumptions of Theorem 134 hold. If $\{p_j\}_{j=1}^M$ is basis for P_θ then the representation*

$$W_{G,X} = \left\{ \sum_{i=1}^{N} \alpha_i G\left(\cdot - x^{(i)}\right) + \sum_{j=1}^{M} \beta_j p_j : P_X^T \alpha = 0, \ \alpha = (\alpha_i), \ \alpha_i, \beta_j \in \mathbb{C} \right\},$$

is unique in terms of α_i and β_j.

Proof. This follows directly from parts 2 and 4 of Theorem 134. ■

In Subsection 4.9.4 a matrix equation will be derived for the coefficients α_i and β_j. From Theorem 47 we know that the data-translated basis functions $\left\{G\left(\cdot - x^{(i)}\right)\right\}_{i=1}^N$ are linearly independent and from the last theorem we know that the dimension of $\dot{W}_{G,X}$ is $N - M$. However no subset of the functions $\left\{G\left(\cdot - x^{(i)}\right)\right\}_{i=1}^N$ containing $N - M$ of these functions spans $\dot{W}_{G,X}$. The functions $\left\{G\left(\cdot - x^{(i)}\right)\right\}_{i=1}^N$ can be called basis functions w.r.t. $\dot{W}_{G,X}$ in the sense that linear combinations of these functions can be used to construct a basis for $\dot{W}_{G,X}$. In fact, in the next corollary we will show that provided the minimal unisolvent subset A used to construct the Riesz representer R_x lies in X and constitutes the first M points of X, it will follow that $\{R_{x^{(i)}}\}_{i=M+1}^N$ is a basis for $\dot{W}_{G,X}$.

Theorem 136 *Suppose w is a weight function with properties W2 and W3 for order θ. Next let $X = \left\{x^{(k)}\right\}_{k=1}^N$ be θ-unisolvent and suppose $X_1 = \left\{x^{(k)}\right\}_{k=1}^M$ is a minimal unisolvent set. Define the Riesz representer function R_x by 4.32 using X_1.*

Then the spaces $W_{G,X}$ and $\dot{W}_{G,X}$ have the following properties:

1. *$W_{G,X}$ has basis $\{R_{x^{(i)}}\}_{i=1}^N$ over the complex numbers.*

2. *$\dot{W}_{G,X}$ has basis $\{R_{x^{(i)}}\}_{i=M+1}^N$ over the complex numbers.*

Proof. It will first be shown that if $\beta = \{\beta_k\}_{k=M+1}^N$ are any complex numbers then $\sum\limits_{k=M+1}^{N} \beta_k R_{x^{(k)}} \in W_{G,X}$. Let $g_\beta = \sum\limits_{k=M+1}^{N} \beta_k R_{x^{(k)}}$. Then each $R_{x^{(k)}} \in X_w^\theta$ and from equation 4.39

$$\left(R_{x^{(k)}}\right)_F = (2\pi)^{-d/2} \frac{\left(\mathcal{Q}_x\left(e^{-i(x,\cdot)}\right)\right)\left(x = x^{(k)}\right)}{w\,|\cdot|^{2\theta}} \ on \ \mathbb{R}^d \backslash 0.$$

Thus

$$\begin{aligned}
(g_\beta)_F &= \sum_{k=M+1}^{N} \beta_k \left(R_{x^{(k)}}\right)_F \\
&= \frac{(2\pi)^{-d/2}}{w\,|\cdot|^{2\theta}} \sum_{k=M+1}^{N} \beta_k \left(\mathcal{Q}_x\left(e^{-i(x,\cdot)}\right)\right)\left(x = x^{(k)}\right) \\
&= \frac{(2\pi)^{-d/2}}{w\,|\cdot|^{2\theta}} \sum_{k=M+1}^{N} \beta_k \left(e^{-i(x,\cdot)} - \sum_{j=1}^{M} l_j(x)\, e^{-i\left(x^{(j)},\cdot\right)}\right)\left(x = x^{(k)}\right) \\
&= \frac{(2\pi)^{-d/2}}{w\,|\cdot|^{2\theta}} \sum_{k=M+1}^{N} \beta_k \left(e^{-i\left(x^{(k)},\cdot\right)} - \sum_{j=1}^{M} l_j\left(x^{(k)}\right) e^{-i\left(x^{(j)},\cdot\right)}\right) \\
&= \frac{(2\pi)^{-d/2}}{w\,|\cdot|^{2\theta}} \left(\sum_{k=M+1}^{N} \beta_k e^{-i\left(x^{(k)},\cdot\right)} - \sum_{j=1}^{M}\left(\sum_{k=M+1}^{N} \beta_k l_j\left(x^{(k)}\right)\right) e^{-i\left(x^{(j)},\cdot\right)}\right).
\end{aligned}$$

From 4.46 and 4.44 we see that if $f_\gamma = \sum_{k=1}^{N} \gamma_k G\left(\cdot - x^{(k)}\right)$ then $(f_\gamma)_F = \frac{(2\pi)^{-d/2}}{w|\cdot|^{2\theta}} \sum_{k=1}^{N} \gamma_k e^{-i\left(x^{(k)},\cdot\right)}$ and setting $(g_\beta)_F = \frac{(2\pi)^{-d/2}}{w|\cdot|^{2\theta}} \sum_{k=1}^{N} \gamma_k e^{-i\left(x^{(k)},\cdot\right)}$ yields

$$\gamma_j = \begin{cases} -\,(2\pi)^{d/2} \sum\limits_{k=M+1}^{N} l_j\left(x^{(k)}\right)\beta_j, & j \le M, \\ (2\pi)^{d/2}\,\beta_j, & j > M. \end{cases} \tag{4.49}$$

The criterion of part 1 of Theorem 105 clearly implies that $P_X^T\gamma = 0$ and thus $f_\gamma = (2\pi)^{-d/2} \sum_{j=1}^{N} \gamma_j G\left(x - x^{(j)}\right) \in \dot{W}_{G,X}$ and $(g_\beta - f_\gamma)_F = 0$. But by part 4 Summary 111, $g_\beta - f_\gamma \in P_\theta$ and thus $g_\beta \in W_{G,X}$. Further, $g_\beta = \sum_{k=M+1}^{N} \beta_k R_{x^{(k)}}$ so $\sum_{k=1}^{N} \beta_k R_{x^{(k)}} \in W_{G,X}$ and $\operatorname{span}\{R_{x^{(j)}}\}_{j=1}^{N} \subset W_{G,X}$.

To prove the converse choose $u \in W_{G,X}$. Then $u = \sum_{j=1}^{N} \beta_j G\left(\cdot - x^{(j)}\right) + q$ for some $q \in P_\theta$ and $\beta \in \mathbb{C}^N$ such that $P_X^T\beta = 0$. When $\gamma = 0$ and $x = x^{(j)}$ equation 4.33 for R_x can be rearranged to give

$$G\left(\cdot - x^{(j)}\right) = (2\pi)^{d/2} R_{x^{(j)}} + \sum_{i=1}^{M} l_i\left(x^{(j)}\right) G\left(\cdot - x^{(i)}\right) - q_{x^{(j)}},$$

where each $q_{x^{(j)}} \in P_\theta$. Hence, since 4.34 implies $q_{x^{(j)}} = (2\pi)^{\frac{d}{2}} R_{x^{(j)}}$ for $j \le M$,

$$\begin{aligned}
\sum_{j=1}^{N} \beta_j G\left(\cdot - x^{(j)}\right) &= (2\pi)^{\frac{d}{2}} \sum_{j=1}^{N} \beta_j R_{x^{(j)}} + \sum_{j=1}^{N} \beta_j \sum_{i=1}^{M} l_i\left(x^{(j)}\right) G\left(\cdot - x^{(i)}\right) - \sum_{j=1}^{N} \beta_j q_{x^{(j)}} \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=M+1}^{N} \beta_j R_{x^{(j)}} + \sum_{i=1}^{M}\left(\sum_{j=1}^{N} \beta_j l_i\left(x^{(j)}\right)\right) G\left(\cdot - x^{(i)}\right) - \sum_{j=M+1}^{N} \beta_j q_{x^{(j)}},
\end{aligned}$$

and so

$$\begin{aligned}
u &= \sum_{j=1}^{N} \beta_j G\left(\cdot - x^{(j)}\right) + q \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=M+1}^{N} \beta_j R_{x^{(j)}} + \sum_{i=1}^{M} \left(\sum_{j=1}^{N} \beta_j l_i(x^{(j)}) \right) G\left(\cdot - x^{(i)}\right) - \sum_{j=M+1}^{N} \beta_j q_{x^{(j)}} + q \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=M+1}^{N} \beta_j R_{x^{(j)}} + \sum_{i=1}^{M} \left(\beta_i + \sum_{j=M+1}^{N} \beta_j l_i(x^{(j)}) \right) G\left(\cdot - x^{(i)}\right) - \sum_{j=M+1}^{N} \beta_j q_{x^{(j)}} + q \\
&= (2\pi)^{\frac{d}{2}} \sum_{j=M+1}^{N} \beta_j R_{x^{(j)}} - \sum_{j=M+1}^{N} \beta_j q_{x^{(j)}} + q,
\end{aligned}$$

since by part 1 Theorem 105, $P_X^T \beta = 0$ iff $\beta_i = -\sum\limits_{j=M+1}^{N} \beta_j l_i\left(x^{(j)}\right)$ for each l_i. Thus $W_{G,X} = \operatorname*{span}_{\mathbb{C}} \{R_{x^{(j)}}\}_{j=1}^{N}$. Further, by part 2 of Theorem 124, we know that $R_{x^{(j)}} = l_j$ for $1 \le j \le M$ and so $\operatorname*{span}_{\mathbb{C}} \{R_{x^{(j)}}\}_{j=1}^{M} = P_\theta$. Therefore the result $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$ from Theorem 134 implies that $\dot{W}_{G,X} = \operatorname*{span}_{\mathbb{C}} \{R_{x^{(j)}}\}_{j=M+1}^{N}$. ∎

The previous theorem assumed that the Riesz representers were constructed from a minimal unisolvent subset X_1 of X which consisted of the first M points of X. The next corollary weakens this assumption by that we can choose any minimal unisolvent subset of X. Now recall Definition 100 regarding permutation theory.

Corollary 137 *Suppose w is a weight function with properties W2 and W3 for order θ. Next let $X = \left\{x^{(k)}\right\}_{k=1}^{N}$ be θ-unisolvent and suppose $A \subset X$ is any minimal unisolvent subset. Denote its cardinal basis by $\{l_k\}_{k=1}^{M}$ and define the Riesz representer function R_x by 4.35. Then the spaces $W_{G,X}$ and $\dot{W}_{G,X}$ have the following properties:*

1. *$W_{G,X}$ has basis $\{R_{x^{(i)}}\}_{i=1}^{N}$ over the complex numbers.*
2. *$\dot{W}_{G,X}$ has basis $\left\{R_{x^{(i)}} : x^{(i)} \notin A\right\}$ over the complex numbers.*

Proof. Part 1 We want to use Theorem 136 so start by re-ordering X using a permutation π such that the first M points in πX belong to A. By Theorem 136, $W_{G,\pi X}$ has basis $\left\{R_{y^{(i)}}(\pi A)\right\}_{i=1}^{N}$ where $y^{(i)} = x^{(\pi(i))}$. By 4.35 the function R_x only depends on G, A and x and by Theorem 119, R_x is independent of the order of A. Hence $W_{G,\pi X}$ has bases $\left\{R_{y^{(i)}}(\pi A)\right\}_{i=1}^{N} = \left\{R_{y^{(i)}}(A)\right\}_{i=1}^{N} = \{R_{x^{(i)}}\}_{i=1}^{N}$. By Theorem 132 the set $W_{G,X}$ is independent of the order of the points in X. Hence $W_{G,\pi X} = W_{G,X}$ and we have part 1.

Part 2 We know from part 4 Theorem 134 that $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$, and from part 2 Theorem 124 we know that $\left\{R_{x^{(i)}} : x^{(i)} \in A\right\}$ is the cardinal basis for P_θ generated by A. Thus $\dot{W}_{G,X}$has basis $\left\{R_{x^{(i)}} : x^{(i)} \notin A\right\}$. ∎

4.8 The vector-valued evaluation operator $\widetilde{\mathcal{E}}_X$

The vector-valued evaluation operator $\widetilde{\mathcal{E}}_X$ and its adjoint will be fundamental to the study of the interpolant in this document and to the smoothers in the later chapters. This function evaluates a function on an ordered set of points to form a complex vector.

Definition 138 ***The vector-valued evaluation operator $\widetilde{\mathcal{E}}_X$***

Let $X = \{x^{(i)}\}_{i=1}^{N}$ *be a set of* N *distinct points in* $\mathbb{R}^d$*. Let* u *be a continuous function. Then the evaluation operator* $\widetilde{\mathcal{E}}_X$ *is defined by*

$$\widetilde{\mathcal{E}}_X u = \left(u\left(x^{(i)}\right)\right)_{i=1}^{N}.$$

Sometimes we will use the more compact notation u_X *for* $\widetilde{\mathcal{E}}_X u$ *and when dealing with matrices* $\widetilde{\mathcal{E}}_X u$ *will be regarded as a column vector.*

We will now assume that the weight function w satisfies properties W2 and W3 for order θ and κ. Then the functions in X_w^θ are continuous. The Riesz representer R_x of the evaluation functional $f \to f(z)$ allows some important properties of the evaluation operator $\widetilde{\mathcal{E}}_X$ to be proved. In the next theorem we assume that the first M points in X are minimally unisolvent and use these to generate R_x.

Theorem 139 *Suppose that* $X = \{x^{(i)}\}_{i=1}^{N}$ *is a* θ*-unisolvent set of distinct points in* $\mathbb{R}^d$*. Assume that* $A = \{a_i\}_{i=1}^{M}$ *is any minimal unisolvent subset with cardinal basis* $\{l_k\}_{k=1}^{M}$ *which we use to define the Riesz representer* R_x*, the Light norm* $\|\cdot\|_{w,\theta}$ *and the Lagrangian interpolation operator* $\mathcal{P}$*. Then:*

1. *The evaluation operator* $\widetilde{\mathcal{E}}_X : \left(X_w^\theta, \|\cdot\|_{w,\theta}\right) \to \left(\mathbb{C}^N, |\cdot|\right)$ *is continuous and onto. Also,* $\operatorname{null}\widetilde{\mathcal{E}}_X = W_{G,X}^{\perp}$.

2. *The adjoint operator* $\widetilde{\mathcal{E}}_X^* : \mathbb{C}^N \to X_w^\theta$*, defined by* $\left(\widetilde{\mathcal{E}}_X f, \beta\right)_{\mathbb{C}^N} = \left(f, \widetilde{\mathcal{E}}_X^* \beta\right)_{w,\theta}$*, satisfies*

$$\widetilde{\mathcal{E}}_X^* \beta = \sum_{i=1}^{N} \beta_i R_{x^{(i)}}, \quad \beta \in \mathbb{C}^N, \tag{4.50}$$

   and $\widetilde{\mathcal{E}}_X^*$ *is continuous, 1-1 and maps* $\left(\mathbb{C}^N, |\cdot|\right)$ *onto* $\left(W_{G,X}, \|\cdot\|_{w,\theta}\right)$.

3. $\left\|\widetilde{\mathcal{E}}_X^*\right\| = \left\|\widetilde{\mathcal{E}}_X\right\| = \|R_{X,X}\|$*, where* $R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right)$ *is the (regular) reproducing kernel matrix discussed in Subsection 4.6.2 and* $\|\cdot\|$ *is the matrix norm corresponding to the Euclidean vector norm.*

4. *Assuming* β *is a column vector and* $\widetilde{\mathcal{E}}_X$ *is a column vector we have*

$$\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* \beta = R_{X,X}\beta, \quad \beta \in \mathbb{C}^N.$$

5. *Assuming that* $\widetilde{\mathcal{E}}_X$ *and* γ *are column vectors,*

$$\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_A^* \gamma = L_X \gamma, \quad \gamma \in \mathbb{C}^M,$$

   where $L_X = \left(l_j\left(x^{(i)}\right)\right)$ *is the* $N \times M$ *cardinal unisolvent matrix (Definition 103) corresponding to the minimal unisolvent set* A.

6. $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X : X_w^\theta \to W_{G,X}$ *is onto and* $\operatorname{null}\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X = W_{G,X}^{\perp}$*. Also, for* $u \in X_w^\theta$

$$\left(\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X u\right)(x) = \sum_{i=1}^{N} u\left(x^{(i)}\right) R_{x^{(i)}}(x) = \left(u, \sum_{j=1}^{N} R_{x^{(j)}}(x)\, R_{x^{(j)}}\right)_{w,\theta}.$$

7. *For* $\alpha \in \mathbb{C}^N$, $\left(\mathcal{P}\widetilde{\mathcal{E}}_X^* \alpha\right)(x) = \left(\alpha^T L_X\right) \widetilde{l}(x)$ *where* $\widetilde{l}(x) = \left(l_i(x)\right)$.

8. $\widetilde{\mathcal{E}}_X \mathcal{P} f = L_X \widetilde{\mathcal{E}}_A f$.

Proof. Parts 1 and 2 We will first show that $\widetilde{\mathcal{E}}_X$ is continuous**.** In fact

$$\begin{aligned}\left|\widetilde{\mathcal{E}}_X u\right|_{\mathbb{C}^N}^2 = \sum_{i=1}^{N}\left|u(x^{(i)})\right|^2 = \sum_{i=1}^{N}\left|(u, R_{x^{(i)}})_{w,\theta}\right|^2 &\le \sum_{i=1}^{N} \|u\|_{w,\theta}^2 \|R_{x^{(i)}}\|_{w,\theta}^2 \\ &= \left(\sum_{i=1}^{N} \|R_{x^{(i)}}\|_{w,\theta}^2\right) \|u\|_{w,\theta}^2,\end{aligned}$$

so that $\widetilde{\mathcal{E}}_X : X_w^\theta \to \mathbb{C}^N$ is continuous. Next we show that $\widetilde{\mathcal{E}}_X$ is onto. The Hilbert space adjoint $\widetilde{\mathcal{E}}_X^*$ of $\widetilde{\mathcal{E}}_X$ is defined by

$$\left(\widetilde{\mathcal{E}}_X u, \beta\right)_{\mathbb{C}^N} = \left(u, \widetilde{\mathcal{E}}_X^* \beta\right)_{w,\theta}. \tag{4.51}$$

We now calculate the adjoint by using the representer R_x to good effect.

$$\left(\widetilde{\mathcal{E}}_X u, \beta\right)_{\mathbb{C}^N} = \sum_{i=1}^N u\left(x^{(i)}\right)\overline{\beta_i} = \sum_{i=1}^N \left(u, R_{x^{(i)}}\right)_{w,\theta}\overline{\beta_i} = \left(u, \sum_{i=1}^N \beta_i R_{x^{(i)}}\right)_{w,\theta},$$

so that

$$\widetilde{\mathcal{E}}_X^* \beta = \sum_{i=1}^N \beta_i R_{x^{(i)}}, \quad \beta \in \mathbb{C}^N. \tag{4.52}$$

In Corollary 137 it was shown that the functions $\{R_{x^{(i)}}\}_{i=1}^N$ form a basis for $W_{G,X}$. Hence $\operatorname{range} \widetilde{\mathcal{E}}_X^* = W_{G,X}$ and $\operatorname{null} \widetilde{\mathcal{E}}_X^* = \{0\}$. We now recall the closed-range theorem e.g. Yosida [24], which states that for a continuous linear operator $\mathcal{S}$ the $\operatorname{range} \mathcal{S}$ is closed iff $\operatorname{range} \mathcal{S}^*$ is closed. Since $\operatorname{range} \widetilde{\mathcal{E}}_X^* = W_{G,X}$ is finite dimensional it is closed so we conclude that $\operatorname{range} \widetilde{\mathcal{E}}_X$ is closed. Consequently, using the result that $\overline{\operatorname{range} \mathcal{S}} = (\operatorname{null} \mathcal{S}^*)^\perp$ for any continuous linear operator $\mathcal{S}$, we have

$$\operatorname{range} \widetilde{\mathcal{E}}_X = \overline{\operatorname{range}\left(\widetilde{\mathcal{E}}_X\right)} = \left(\operatorname{null} \widetilde{\mathcal{E}}_X^*\right)^\perp = \{0\}^\perp = \mathbb{R}^N.$$

Finally, from 4.51, $\widetilde{\mathcal{E}}_X u = 0$ iff $\left(\widetilde{\mathcal{E}}_X u, \beta\right)_{\mathbb{C}^N} = 0$ for all $\beta \in \mathbb{C}^N$ iff $\left(u, \widetilde{\mathcal{E}}_X^* \beta\right)_{w,\theta} = 0$ for all $\beta \in \mathbb{C}^N$. But $\operatorname{range} \widetilde{\mathcal{E}}_X^* = W_{G,X}$ so $\widetilde{\mathcal{E}}_X u = 0$ iff $(u, f)_{w,\theta} = 0$ for all $f \in W_{G,X}$ i.e. iff $u \in W_{G,X}^\perp$.

Part 3 That $\left\|\widetilde{\mathcal{E}}_X^*\right\| = \left\|\widetilde{\mathcal{E}}_X\right\|$ is an elementary property of the adjoint. Now

$$\begin{aligned}
\left\|\widetilde{\mathcal{E}}_X^* \beta\right\|_{w,\theta}^2 = \left(\sum_{i=1}^N \beta_i R_{x^{(i)}}, \sum_{j=1}^N \beta_j R_{x^{(j)}}\right)_{w,\theta} &= \sum_{i,j=1}^N \beta_i \overline{\beta_j} \left(R_{x^{(i)}}, R_{x^{(j)}}\right)_{w,\theta} \\
&= \sum_{i,j=1}^N \beta_i \overline{\beta_j} R_{x^{(i)}}\left(x^{(j)}\right) \\
&= \sum_{i,j=1}^N \beta_i \overline{\beta_j} \overline{R_{x^{(j)}}\left(x^{(i)}\right)} = \beta^T \overline{R_{X,X}\beta} \\
&= \overline{\beta}^T R_{X,X}\beta,
\end{aligned}$$

so that, $\left\|\widetilde{\mathcal{E}}_X^*\right\| = \max\limits_{\beta \in \mathbb{C}^N} \frac{\|\widetilde{\mathcal{E}}_X^* \beta\|_{w,\theta}}{|\beta|} = \max\limits_{\beta \in \mathbb{C}^N} \frac{\sqrt{\overline{\beta}^T R_{X,X}\beta}}{|\beta|} = \max\limits_{\beta \in \mathbb{C}^N} \frac{\sqrt{\beta^T R_{X,X}\overline{\beta}}}{|\beta|}$ and this expression is the largest (positive) eigenvalue of the Hermitian matrix $R_{X,X}$. But this is also the value of $\|R_{X,X}\|$ so $\left\|\widetilde{\mathcal{E}}_X^*\right\| = \|R_{X,X}\|$.

Part 4

$$\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* \beta = \widetilde{\mathcal{E}}_X \left(\sum_{j=1}^N \beta_j R_{x^{(j)}}\right) = \left(\sum_{j=1}^N \beta_j R_{x^{(j)}}\left(x^{(i)}\right)\right)_{j=1,N} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right)\beta = R_{X,X}\beta.$$

Part 5 By part 2 Theorem 124, $R_{a_j} = l_j$ for $j \le M$

$$\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_A^* \gamma = \widetilde{\mathcal{E}}_X \left(\sum_{j=1}^M \gamma_j R_{a_j}\right) = \widetilde{\mathcal{E}}_X \left(\sum_{j=1}^M \gamma_j l_j\right) = \left(\sum_{j=1}^M \gamma_j l_j\left(x^{(i)}\right)\right)_{i=1,N} = L_X \gamma.$$

Part 6 By part 1, $\widetilde{\mathcal{E}}_X : X_w^\theta \to \mathbb{C}^N$ is onto and null $\widetilde{\mathcal{E}}_X = W_{G,X}^\perp$. By part 2, $\widetilde{\mathcal{E}}_X^* : \mathbb{C}^N \to W_{G,X}$ is onto and 1-1. These facts directly imply $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X : X_w^\theta \to W_{G,X}$ is onto and null $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X = W_{G,X}^\perp$.

$$\left(\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f\right)(x) = \left(\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f, R_x\right)_{w,\theta} = \left(\widetilde{\mathcal{E}}_X f, \widetilde{\mathcal{E}}_X R_x\right)_{\mathbb{C}^N} = \sum_{j=1}^N f\left(x^{(j)}\right) \overline{R_x\left(x^{(j)}\right)} = \sum_{j=1}^N f\left(x^{(j)}\right) R_{x^{(j)}}(x),$$

so that

$$\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f = \sum_{j=1}^N f\left(x^{(j)}\right) R_{x^{(j)}}. \tag{4.53}$$

Also $\left(\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f\right)(x) = \left(f, \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X R_x\right)_{w,\theta}$, but by 4.53, $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X R_x = \sum_{j=1}^N R_{x^{(j)}}(x) R_{x^{(j)}}$, so that

$$\left(\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f\right)(x) = \left(f, \sum_{j=1}^N R_{x^{(j)}}(x) R_{x^{(j)}}\right)_{w,\theta}.$$

Part 7 We must show that $\left(\mathcal{P}\widetilde{\mathcal{E}}_X^* \alpha\right)(x) = \alpha^T L_X \widetilde{l}(x)$.

$$\begin{aligned}
\left(\mathcal{P}\widetilde{\mathcal{E}}_X^* \alpha\right)(x) = \mathcal{P}\sum_{j=1}^N \alpha_j R_{x^{(j)}}(x) = \sum_{j=1}^N \alpha_j \mathcal{P} R_{x^{(j)}}(x) &= \sum_{j=1}^N \alpha_j \sum_{i=1}^M R_{x^{(j)}}(a_i)\, l_i(x) \\
&= \sum_{j=1}^N \alpha_j \sum_{i=1}^M R_{a_i}\left(x^{(j)}\right) l_i(x) \\
&= \sum_{j=1}^N \alpha_j \sum_{i=1}^M l_i\left(x^{(j)}\right) l_i(x) \\
&= \alpha^T L_X \widetilde{l}(x).
\end{aligned}$$

Part 8 Since $L_X = \left(l_j\left(x^{(i)}\right)\right)$

$$\widetilde{\mathcal{E}}_X \mathcal{P} f = \widetilde{\mathcal{E}}_X \sum_{j=1}^M f(a_j)\, l_j = \sum_{j=1}^M f(a_j)\, \widetilde{\mathcal{E}}_X l_j = L_X \widetilde{\mathcal{E}}_A f.$$

■

4.9 Variational interpolation with basis functions

The mathematical machinery we have developed above will now be applied to studying two variational interpolation problems. These will be called the minimal norm and minimal seminorm problems respectively. The minimal seminorm problem will involve minimizing the seminorm of the space X_w^θ over the functions in X_w^θ which interpolate the data points $\left\{\left(x^{(i)}, y_i\right)\right\}_{i=1}^N$. The minimal norm problem will involve minimizing the Light norm of the space X_w^θ over the functions in X_w^θ which interpolate the data. We show the minimal norm problem has a unique solution and then show the minimal seminorm problem has the same, unique solution and that this solution lies in the finite dimensional space $W_{G,X}$ where G is an order θ basis function of the weight function w and $X = \{x^{(i)}\}_{i=1}^N$ i.e. the solution is a basis function interpolant. A matrix equation is also derived for the coefficients of the basis functions.

To render the interpolation problem meaningful we suppose the weight function w has weight properties W2 and W3 for order θ and smoothness parameter κ, so that X_w^θ is a space of continuous functions. Further, the conditions on w allow us to define a continuous basis function G of order θ.

4.9.1 The minimal seminorm interpolation problem

The scattered data set to be interpolated consists of interpolating the N data points

$$\left\{\left(x^{(i)}, y_i\right)\right\}_{i=1}^{N}, \quad x^{(i)} \in \mathbb{R}^d,\ x^{(i)}\ \mathit{distinct},\ y_i \in \mathbb{R}.$$

We call $X = \{x^{(i)}\}_{i=1}^{N}$ the independent data and $y = \{y_i\}_{i=1}^{N}$ the dependent data, and the data will sometimes be referred to as $[X, y]$. The points in X are not completely unconstrained because it will be assumed that X is unisolvent with respect to the polynomials P_θ (Definition 99). We now require that a *minimal seminorm interpolant* u_I satisfies

$$|u_I|_{w,\theta} = \min\left\{|u|_{w,\theta} : u \in X_w^\theta;\ u_I\left(x^{(i)}\right) = y_i,\ i = 1, \ldots, N\right\}.$$

4.9.2 The minimal norm interpolation problem

Since the independent data X is θ-unisolvent, from Definition 99 there is at least one minimal unisolvent subset of X, say A. Denote by b the dependent data corresponding to A. The Light norm $\|\cdot\|_{w,\theta}$ of 4.29 is now constructed using A. The scattered data is the same as for the minimal seminorm interpolation problem i.e. the *minimal norm interpolant* u_I satisfies

$$\|u_I\|_{w,\theta} = \min\left\{\|u_I\|_{w,\theta} : u \in X_w^\theta,\ u_I\left(x^{(i)}\right) = y_i\ \mathit{for}\ i = 1, \ldots, N\right\}. \tag{4.54}$$

Using Hilbert space techniques, we now want to do the following:

1. Show their exists a unique minimal norm interpolant.
2. Show their exists a unique minimal seminorm interpolant.
3. Show these interpolants are identical.
4. Show the interpolant is a basis function interpolant lying in the space $W_{G,X}$ i.e. it can be written as a linear combination of data-translated basis functions plus a polynomial of degree less than θ.

4.9.3 Solving the minimal interpolation problems

In this subsection we show the minimal norm interpolation problem has a unique solution which we call the minimal norm interpolant. It is then shown that this solution is also the unique solution to the minimum seminorm interpolation problem. We then derive a formula for this solution which implies that the solution lies in the finite dimensional space $W_{G,X}$ introduced in Definition 130. The last result will allow us to derive matrix results for the interpolant in the next subsection.

Theorem 140 *Endow X_w^θ with the Light norm $\|\cdot\|_{w,\theta}$ defined by 4.29 using A. Then there exists a unique minimal norm interpolant. In fact, given $y \in \mathbb{R}^N$ there is a unique interpolant $u_I \in X_w^\theta$ such that*

$$\|u_I\|_{w,\theta} = \min\left\{\|u\|_{w,\theta} : u \in X_w^\theta\ \mathit{and}\ \widetilde{\mathcal{E}}_X u = y\right\}. \tag{4.55}$$

If $v \in X_w^\theta$ is any other interpolant satisfying $\widetilde{\mathcal{E}}_X v = y$ then

$$\|u_I\|_{w,\theta}^2 + \|v - u_I\|_{w,\theta}^2 = \|v\|_{w,\theta}^2, \tag{4.56}$$

or equivalently

$$(v - u_I, u_I)_{w,\theta} = 0. \tag{4.57}$$

Proof. Now by part 2 of Theorem 139 the evaluation operator $\widetilde{\mathcal{E}}_X : X_w^\theta \to \mathbb{R}^N$ is continuous and onto, and since the singleton set $\{y\}$ is closed, it follows that the set

$$\left\{u : \widetilde{\mathcal{E}}_X u = y\right\} = \widetilde{\mathcal{E}}_X^{-1}(y),$$

is a non-empty, proper, closed subspace of the Hilbert space X_w^θ. Hence this subspace contains a unique element of smallest norm, say u_I, which satisfies equation 4.55. If $\widetilde{\mathcal{E}}_X v = y$ then

$$\left\{u : u \in X_w^\theta \text{ and } \widetilde{\mathcal{E}}_X u = y\right\} = \left\{v - s : s \in \operatorname{null} \widetilde{\mathcal{E}}_X\right\}.$$

Now

$$\min\left\{\|v - s\|_{w,\theta} : s \in \operatorname{null} \widetilde{\mathcal{E}}_X\right\} = \operatorname{dist}\left(v, \operatorname{null} \widetilde{\mathcal{E}}_X\right),$$

is the distance between v and the closed subspace $\operatorname{null} \widetilde{\mathcal{E}}_X$. Therefore there exists a unique $s_I \in \operatorname{null} \widetilde{\mathcal{E}}_X$ such that

$$u_I = v - s_I, \tag{4.58}$$

$$\|v - s_I\|_{w,\theta} = \min\left\{\|v - s\|_{w,\theta} : s \in \operatorname{null} \widetilde{\mathcal{E}}_X\right\},$$

and

$$\|s_I\|_{w,\theta}^2 + \|v - s_I\|_{w,\theta}^2 = \|v\|_{w,\theta}^2. \tag{4.59}$$

Substituting for s_I in 4.59 using 4.58 yields 4.56. Equation 4.57 follows since it is a necessary and sufficient condition for 4.56 to be true. ■

Theorem 141 *The unique minimal norm interpolant u_I of Theorem 140 is the unique minimal semi-norm interpolant. In fact, suppose $y \in \mathbb{R}^N$. Then*

$$|u_I|_{w,\theta} = \min\left\{|u|_{w,\theta} : u \in X_w^\theta \textit{ and } \widetilde{\mathcal{E}}_X u = y\right\}. \tag{4.60}$$

If v is any other interpolant satisfying $\widetilde{\mathcal{E}}_X v = y$ then

$$|u_I|_{w,\theta}^2 + |v - u_I|_{w,\theta}^2 = |v|_{w,\theta}^2, \tag{4.61}$$

or equivalently

$$\langle v - u_I, u_I\rangle_{w,\theta} = 0. \tag{4.62}$$

Proof. Recall that the Light norm 4.29 for the minimal norm interpolant problem 4.54 was constructed using the minimally unisolvent subset A of independent data which had corresponding dependent data y'. Since $\widetilde{\mathcal{E}}_X f = y$

$$\|f\|_{w,\theta}^2 = |f|_{w,\theta}^2 + \sum_{x^{(i)} \in A} \left|f(x^{(i)})\right|^2 = |f|_{w,\theta}^2 + |y'|^2,$$

By Theorem 140 there exists a unique minimal norm interpolant u_I such that

$$\|u_I\|_{w,\theta} = \min\left\{\|u\|_{w,\theta} : u \in X_w^\theta \textit{ and } \widetilde{\mathcal{E}}_X u = y\right\},$$

and hence

$$\begin{aligned} |u_I|_{w,\theta}^2 + |b|^2 &= \min\left\{|u|_{w,\theta}^2 + |y'|^2 : u \in X_w^\theta \textit{ and } \widetilde{\mathcal{E}}_X u = y\right\} \\ &= \min\left\{|u|_{w,\theta}^2 : u \in X_w^\theta \textit{ and } \widetilde{\mathcal{E}}_X u = y\right\} + |y'|^2, \end{aligned}$$

so that

$$\|u_I\|_{w,\theta} = \min\left\{|u|_{w,\theta} : u \in X_w^\theta \textit{ and } \widetilde{\mathcal{E}}_X u = y\right\}.$$

To prove equalities 4.61 and 4.62 first note that if $\widetilde{\mathcal{E}}_X f = 0$ then $\|f\|_{w,\theta}^2 = |f|_{w,\theta}^2 + \sum_{i=1}^M \left|f(x^{(i)})\right|^2 = |f|_{w,\theta}^2$. Thus, since $\widetilde{\mathcal{E}}_X (v - u_I) = 0$,

$$\begin{aligned} |u_I|_{w,\theta}^2 + |v - u_I|_{w,\theta}^2 = |u_I|_{w,\theta}^2 + \|v - u_I\|_{w,\theta}^2 &= \|u_I\|_{w,\theta}^2 - |y'|^2 + \|v - u_I\|_{w,\theta}^2 \\ &= \|v\|_{w,\theta}^2 - |y'|^2, \quad \textit{by } 4.56, \\ &= |v|_{w,\theta}^2, \end{aligned}$$

which proves 4.61. To prove 4.62 we use equation 4.57 to obtain:

$$0 = (v - u_I, u_I)_{w,\theta} = \langle v - u_I, u_I \rangle_{w,\theta} + \sum_{i=1}^{M} (v - u_I)\left(x^{(i)}\right) \overline{u_I\left(x^{(i)}\right)} = \langle v - u_I, u_I \rangle_{w,\theta},$$

since $v\left(x^{(i)}\right) = u_I\left(x^{(i)}\right)$ for $i = 1, \ldots, N$. ■

There is a very close connection between the minimal norm interpolants, the space $W_{G,X}$ and the adjoint $\widetilde{\mathcal{E}}_X^*$ 4.52 of the evaluation operator $\widetilde{\mathcal{E}}_X$. The next theorem describes this connection and characterizes the space of minimal norm interpolants associated with a given unisolvent set of independent data points X.

The minimal norm interpolation problem was formulated using a Light norm $\|\cdot\|_{w,\theta}$ 4.29 constructed with a minimal unisolvent set $A \subset X$ and the corresponding cardinal basis of A. In the next theorem we will also construct the Riesz representer R_x 4.35 using A and its cardinal basis. Then by Definition 128 the reproducing kernel matrix $R_{X,X}$ is given by $R_{X,X} = \left(R_{x^{(j)}}(x^{(i)})\right)$.

Theorem 142 *We know from Theorem 140 that for given dependent data $[X; y]$ we can associate a unique minimal norm interpolant $u_I \in X_w^\theta$. In fact*

$$u_I = \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y \in W_{G,X}, \tag{4.63}$$

and the space of all minimal norm interpolants is $W_{G,X}$. Here $R_{X,X}$ is the reproducing kernel matrix.

Proof. From parts 3 and 4 of Theorem 139 $\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* = R_{X,X}$ and so

$$\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* y = \left(\left(\widetilde{\mathcal{E}}_X^* y\right)\left(x^{(i)}\right)\right) = R_{X,X} y,$$

where $R_{X,X}$ is regular by part 2 Theorem 129. In general $R_{X,X} \neq I$ so in general $\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* y \neq y$ and $\widetilde{\mathcal{E}}_X^* y$ is not an interpolant of y.

However, $\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} = I$, so $\widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y$ is an interpolant and
$\widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y \in W_{G,X}$. For convenience set $u = \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y$. We want to show that $u = u_I$. But since both u_I and u interpolate y

$$\begin{aligned}(u, u_I - u)_{w,\theta} = \left(\widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y, u_I - u\right)_{w,\theta} &= \left(\left(R_{X,X}\right)^{-1} y, \widetilde{\mathcal{E}}_X \left(u_I - u\right)\right)_{w,\theta} \\ &= 0,\end{aligned}$$

and so

$$\|u_I\|_{w,\theta}^2 = \|u_I - u + u\|_{w,\theta}^2 = \|u_I - u\|_{w,\theta}^2 + \|u\|_{w,\theta}^2.$$

If $u \neq u_I$ then $\|u\|_{w,\theta} < \|u_I\|_{w,\theta}$, contradicting the fact that u_I is the minimal norm interpolant. Because $R_{X,X}$ is regular, $\left(R_{X,X}\right)^{-1}$ is an isomorphism from $\mathbb{R}^N$ to $\mathbb{R}^N$, and because $\widetilde{\mathcal{E}}_X^*$ is an isomorphism from $\mathbb{R}^N$ to $W_{G,X}$, the formula $u_I = \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y$ implies that range $\left(\widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1}\right) = W_{G,X}$. ■

4.9.4 Matrix equations for the minimal seminorm interpolant

We now know from Theorem 142 that the minimal seminorm interpolant lies in the finite dimensional space $W_{G,X}$. Hence we should now be able to derive a matrix equation for the coefficients of the $G\left(\cdot - x^{(k)}\right)$. This equation will be constructed using the basis function matrix and the unisolvent matrix. This will require the following lemma which turns out to be very useful since it provides key results for a block matrix which has a structure central to the study of the matrix equations in both basis function interpolation and basis function smoothing theory.

Lemma 143 *Let B be a complex-valued matrix and C be a real-valued matrix. Suppose the block matrix $\begin{pmatrix} B & C \\ C^T & O \end{pmatrix}$ is square.*

Suppose further that for complex vectors z

$$z^T B \overline{z} = 0 \text{ and } C^T z = 0 \text{ implies } z = 0. \tag{4.64}$$

1. *Then the equation*
$$\begin{pmatrix} B & C \\ C^T & O \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
implies $u = 0$ *and* $v \in \operatorname{null} C$.

2. *If, further to part 1,* $\operatorname{null} C = \{0\}$ *then the block matrix is regular.*

Proof. Part 1 We proceed by showing that
$$\begin{pmatrix} B & C \\ C^T & O \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
implies $u = 0$. Expanding this block matrix equation yields $Bu + Cv = 0$ and $C^T u = 0$. Thus
$$0 = u^T (\overline{B}\overline{u} + \overline{C}\overline{v}) = u^T \overline{B}\overline{u} + u^T \overline{C}\overline{v} = u^T \overline{B}\overline{u} + (C^T u)^T \overline{v} = u^T \overline{B}\overline{u},$$
and so $\overline{u}^T Bu = 0$. Set $x = \overline{u}$. Then $x^T B\overline{x} = 0$ and $C^T x = 0$, and so 4.64 implies $x = 0$. Hence $u = 0$ and $Bv = 0$.

Part 2 Clearly if $\operatorname{null} C = \{0\}$, by part 1 the block matrix has null space zero and so is regular. ■

We now derive the matrix equation for minimal seminorm interpolant:

Theorem 144 *Suppose* $\{p_j\}_{j=1}^M$ *is any basis for* P_θ. *The space* $W_{G,X}$ *contains only one minimal seminorm interpolant* u_I *to any given data vector* $y \in \mathbb{R}^N$. *This minimal seminorm interpolant* u_I *is given uniquely by*
$$u_I(x) = \sum_{i=1}^N v_i G\left(x - x^{(i)}\right) + \sum_{j=1}^M \beta_j p_j(x), \tag{4.65}$$
where the coefficient vectors $v = (v_i)$ *and* $\beta = \left(\beta_j\right)$ *satisfy the matrix equation*
$$\begin{pmatrix} G_{X,X} & P_X \\ P_X^T & O \end{pmatrix} \begin{pmatrix} v \\ \beta \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}. \tag{4.66}$$
Here P_X *is the unisolvency matrix* $\left(p_j\left(x^{(i)}\right)\right)$ *and* $G_{X,X}$ *is the basis function matrix* $\left(G\left(x^{(i)} - x^{(j)}\right)\right)$. *Also, the matrix* $\begin{pmatrix} G_{X,X} & P_X \\ P_X^T & O \end{pmatrix}$ *has size* $(N+M) \times (N+M)$ *and is regular.*

Proof. By Theorem 142, $W_{G,X}$ contains only one minimal seminorm interpolant. Using equation 4.65 and the interpolation requirement $u_I\left(x^{(i)}\right) = y_i$ we obtain the pair of equations
$$G_{X,X} v + P_X \beta = y; \quad P_X^T v = 0,$$
or in block matrix form
$$\begin{pmatrix} G_{X,X} & P_X \\ P_X^T & O \end{pmatrix} \begin{pmatrix} v \\ \beta \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}.$$
By part 3 Theorem 134, $v^T G_{X,X} \overline{v} = 0$ and $P_X^T v = 0$ implies $v = 0$, and since X is unisolvent, part 1 Theorem 104 implies $\operatorname{null} P_X = \{0\}$. Thus, by Lemma 143 above, the matrix of 4.66 is regular and the interpolation problem has a unique solution $\left(v^T, \beta^T\right)$. ■

From part 1 Corollary 137 $\{R_{x^{(i)}}\}_{i=1}^N$ is a basis for $W_{G,X}$. In the next theorem we give the corresponding matrix equation for the minimal norm interpolant.

Theorem 145 *We know from Theorem 140 that for given independent data* $X = \left\{x^{(i)}\right\}$ *and dependent data* $y \in \mathbb{R}^N$ *we can associate a unique minimal norm interpolant* $u_I \in X_w^\theta$. *In fact, if* $R_{X,X} = \left\{R_{x^{(j)}}\left(x^{(i)}\right)\right\}$ *is the reproducing kernel matrix then*
$$u_I = \sum_{i=1}^N v_i R_{x^{(i)}}, \quad \textit{when } R_{X,X} v = y, \tag{4.67}$$
where R_x *is the Riesz representer of the evaluation functional* $f \to f(x)$.

Proof. $u_I = \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} y = \widetilde{\mathcal{E}}_X^* v = \sum_{i=1}^N v_i R_{x^{(i)}}$. ■

4.9.5 The minimal interpolant mapping and data functions

The last theorem allows us to define the mapping between a data function and its corresponding interpolant. We call this the *minimal interpolant mapping*:

Definition 146 ***Data functions and the minimal interpolant mapping*** $\mathcal{I}_X : X_w^\theta \to W_{G,X}$
Given an independent data set X, we shall assume that each member of X_w^θ can act as a legitimate data function f and generate the data vector $\widetilde{\mathcal{E}}_X f$.
Equation 4.63 of Theorem 142 enables us to define the linear mapping $\mathcal{I}_X : X_w^\theta \to W_{G,X}$ from the data functions to the corresponding unique minimal interpolant $\mathcal{I}_X f = u_I$ given by

$$\mathcal{I}_X f = \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} \widetilde{\mathcal{E}}_X f, \quad f \in X_w^\theta. \tag{4.68}$$

The purpose of the minimal interpolant mapping will be to study the pointwise convergence of the interpolant to its data function.

Theorem 147 *The minimal interpolant mapping $\mathcal{I}_X : X_w^\theta \to W_{G,X}$ has the following properties:*

1. $\|\mathcal{I}_X f\|_{w,\theta} \le \|f\|_{w,\theta}$ *and* $\|(I - \mathcal{I}_X) f\|_{w,\theta} \le \|f\|_{w,\theta}$.
2. $|\mathcal{I}_X f|_{w,\theta} \le |f|_{w,\theta}$ *and* $|(I - \mathcal{I}_X) f|_{w,\theta} \le |f|_{w,\theta}$.
3. $\mathcal{I}_X$ *is a continuous projection onto* $W_{G,X}$ *with null space* $W_{G,X}^{\perp}$.
4. $\mathcal{I}_X$ *is self-adjoint.*

Proof. Part 1 Follows from 4.56 since the data function is an interpolant.

Part 2 Follows from part 1 since $\mathcal{I}_X$ is an interpolant.

Part 3 Part 1 implies continuity. Part 4 of Theorem 139 implies $\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* = R_{X,X}$ and $R_{X,X}$ is Hermitian and regular. Hence 4.68 implies $\mathcal{I}_X$ is a projection and thus onto. Further, $\widetilde{\mathcal{E}}_X^*$ is 1-1 by part 2 of Theorem 139 so 4.68 implies $\mathcal{I}_X f = 0$ iff $\widetilde{\mathcal{E}}_X f = 0$ iff $f \in W_{G,X}^{\perp}$ by part 1 of Theorem 139.

Part 4 By 4.68, $\left(\mathcal{I}_X f, g\right)_{w,\theta} = \left(\widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} \widetilde{\mathcal{E}}_X f, g\right)_{w,\theta} =$
$= \left(\left(R_{X,X}\right)^{-1} \widetilde{\mathcal{E}}_X f, \widetilde{\mathcal{E}}_X g\right)_{w,\theta} = \left(f, \widetilde{\mathcal{E}}_X^* \left(R_{X,X}\right)^{-1} \widetilde{\mathcal{E}}_X g\right)_{w,\theta}$ since $R_{X,X}$ is Hermitian. ■

4.10 Convergence orders for the minimal interpolant

The estimates for the order of pointwise convergence obtained here are derived using similar techniques to those of Light and Wayne in [9] and [10]. We use the Lagrange interpolation theory from [10] and the function $r_x(x)$.

4.10.1 A unisolvency lemma

To study the convergence of the minimal interpolant we will also need the following lemma which supplies some elementary results from the theory of Lagrange interpolation. These results are stated without proof. This lemma has been created from Lemma 3.2, Lemma 3.5 and the first two paragraphs of the proof of Theorem 3.6 of Light and Wayne [10]. The results of this lemma do not involve any reference to weight or basis functions or functions in X_w^θ, but consider the properties of the set which contains the independent data points and the order of the unisolvency used for the interpolation. Thus we have separated the part of the proof that involves weight functions from the part that uses the detailed theory of Lagrange interpolation operators.

Lemma 148 *Assume first that:*

1. Ω *is a bounded, open, connected subset of* $\mathbb{R}^d$ *having the cone property.*
2. X *is a unisolvent subset of* Ω *of order* θ.

Suppose $\{l_j\}_{j=1}^{M}$ is the cardinal basis of P_θ with respect to a minimal unisolvent set in Ω. Using Lagrange interpolation techniques, it can be shown there exists a constant $K'_{\Omega,\theta} > 0$ such that

$$\sum_{j=1}^{M} |l_j(x)| \le K'_{\Omega,\theta}, \quad x \in \overline{\Omega},$$

and all minimal unisolvent subsets of $\overline{\Omega}$. Now define the data density by

$$h_X = \sup_{\omega \in \Omega} \operatorname{dist}(\omega, X), \tag{4.69}$$

and fix $x \in \Omega$. Using Lagrange interpolation techniques it can be shown there are constants $c_{\Omega,\theta}, h_{\Omega,\theta} > 0$ such that when $h_X < h_{\Omega,\theta}$ there exist a minimal unisolvent set $A \subset X$ satisfying

$$\operatorname{diam} A_x \le c_{\Omega,\theta} h_X,$$

where $A_x = A \cup \{x\}$.

Remark 149 *If ω_1 is one of the furthest points in $\overline{\Omega}$ from X then $h_X = \operatorname{dist}(\omega_1, X)$. h_X can also be interpreted as the radius of the largest sphere centred in $\overline{\Omega}$ that can be placed between the points in X.*

4.10.2 An upper bound for $\sqrt{r_x(x)}$

Theorem 150 *Suppose w is a weight function with properties W2 and W3 for order θ and κ, and that G is a basis function of order θ. Suppose $A = \{a_k\}_{k=1}^{M}$ is a minimal θ-unisolvent set and that $\{l_k\}_{k=1}^{M}$ is the corresponding unique cardinal basis for P_θ. Now construct $\mathcal{P}, \mathcal{Q}$ and R_x using A and $\{l_k\}_{k=1}^{M}$. Then if $r_x = \mathcal{Q}R_x$*

$$\sqrt{r_x(x)} \le \frac{d^{\eta/2}}{(2\pi)^{\frac{d}{4}}\sqrt{(2\eta)!}} \left(1 + \sum_{k=1}^{M} |l_k(x)|\right) (\operatorname{diam} A_x)^\eta \max_{\substack{|\beta| = 2\eta \\ |y| \le \operatorname{diam} A_x}} \left|D^\beta G(y)\right|, \quad x \in \mathbb{R}^d,$$

where $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$ and $A_x = A \cup \{x\}$.

Proof. From part 4 of Theorem 125, $r_x(y) = \langle r_x, r_y \rangle_{w,\theta}$ so that $r_x(x) = |r_x|_{w,\theta}^2 \ge 0$.

Also from Theorem 125, $r_x(y) = (2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x G(y-x)$ when $x \ne y$.

Now from Theorem 50 or 49, $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ and on expanding G about the origin using the Taylor's series with integral remainder in the Appendix A.8 we get

$$G(y-x) = \sum_{|\beta| < 2\eta} \frac{(y-x)^\beta}{\beta!} \left(D^\beta G\right)(0) + \mathcal{R}_{2\eta}(0, y-x). \tag{4.70}$$

To calculate $r_x(y)$ apply the operator $\mathcal{Q}_y \mathcal{Q}_x$ to 4.70 and then, noting that $2\eta \le 2\theta$, use Theorem 118 to eliminate the power series and obtain

$r_x(y) = (2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x \mathcal{R}_{2\eta}(0, y-x)$. Expanding $\mathcal{Q}_y$ and $\mathcal{Q}_x$ using $\mathcal{P}_y$ and $\mathcal{P}_x$ now gives

$$\begin{aligned}
(2\pi)^{\frac{d}{2}} r_x(y) &= \mathcal{R}_{2\eta}(0, y-x) - \mathcal{P}_x(\mathcal{R}_{2\eta}(0, y-x)) - \mathcal{P}_y(\mathcal{R}_{2\eta}(0, y-x)) + \mathcal{P}_y \mathcal{P}_x(\mathcal{R}_{2\eta}(0, y-x)) \\
&= \mathcal{R}_{2\eta}(0, y-x) - \sum_{j=1}^{M} \mathcal{R}_{2\eta}(0, y-a_j) l_j(x) - \sum_{k=1}^{M} \mathcal{R}_{2\eta}(0, a_k - x) l_k(y) + \\
&\quad + \sum_{j,k=1}^{M} \mathcal{R}_{2\eta}(0, a_k - a_j) l_j(x) l_k(y),
\end{aligned}$$

so that

$$\begin{aligned}
(2\pi)^{\frac{d}{2}} r_x(x) &\le \sum_{j=1}^{M} |\mathcal{R}_{2\eta}(0, x-a_j)| \; |l_j(x)| + \sum_{k=1}^{M} |\mathcal{R}_{2\eta}(0, a_k - x)| \; |l_k(x)| + \\
&\quad + \sum_{j,k=1}^{M} |\mathcal{R}_{2\eta}(0, a_k - a_j)| \; |l_j(x)| \; |l_k(x)| \\
&\le 2\left(\sum_{k=1}^{M} |l_k(x)|\right) \max_k |\mathcal{R}_{2\eta}(0, x-a_k)| + \left(\sum_{k=1}^{M} \left|l_{k(x)}\right|\right)^2 \max_{j,k} |\mathcal{R}_{2\eta}(0, a_k - a_j)|.
\end{aligned}$$

The remainder estimate A.4 for the Taylor series in Appendix A.8, namely

$$\left|\mathcal{R}_{2\eta}(0,b)\right| \le \frac{d^{\eta}\,|b|^{2\eta}}{(2\eta)!} \max_{\substack{|\beta|=2\eta \\ y\in[0,b]}} \left|D^{\beta}G(y)\right|,$$

is now applied to the last inequality to obtain

$$\begin{aligned}
(2\pi)^{\frac{d}{2}}\, r_x(x) &\le \frac{d^{\eta/2}}{(2\eta)!} 2\left(\sum_{k=1}^{M} |l_k(x)|\right)\left(\max_k |x-a_k|\right)^{2\eta} \max_{\substack{|\beta|=2\eta \\ |y|\le \operatorname{diam} A_x}} \left|D^{\beta}G(y)\right| + \\
&\quad + \frac{d^{\eta/2}}{(2\eta)!}\left(\sum_{k=1}^{M} |l_k(x)|\right)^2 \left(\max_{j,k} |a_j-a_k|\right)^{2\eta} \max_{\substack{|\beta|=2\eta \\ |y|\le \operatorname{diam} A_x}} \left|D^{\beta}G(y)\right| \\
&< \frac{d^{\eta/2}}{(2\eta)!}\left(1+\sum_{k=1}^{M} |l_k(x)|\right)^2 (\operatorname{diam} A_x)^{2\eta} \max_{\substack{|\beta|=2\eta \\ |y|\le \operatorname{diam} A_x}} \left|D^{\beta}G(y)\right|,
\end{aligned}$$

so that

$$\sqrt{r_x(x)} \le \frac{d^{\eta/2}}{(2\pi)^{\frac{d}{4}}\sqrt{(2\eta)!}}\left(1+\sum_{k=1}^{M} |l_k(x)|\right)(\operatorname{diam} A_x)^{\eta} \max_{\substack{|\beta|=2\eta \\ |y|\le \operatorname{diam} A_x}} \left|D^{\beta}G(y)\right|,$$

as claimed. ∎

4.10.3 The order of pointwise convergence of the interpolant to its data function

The result of the previous subsection now allow us to prove the pointwise convergence of the minimal interpolant to its data function and to derive an order of convergence. The concept of interpolation error and convergence go hand in hand. The interpolation error is simply the difference between the interpolant and the data function at a given point.

Before deriving the interpolation error estimate I will show that there exists a nested sequence of independent data sets with sparsity tending to zero.

Theorem 151 *Suppose Ω is a bounded, open set containing all the independent data sets. Then there exists a sequence of independent data sets $X^{(k)} \subset \Omega$ such that $X^{(k)} \subset X^{(k+1)}$ and $h_{X^{(k)}} \to 0$ as $k \to \infty$.*

Proof. For $k = 1, 2, 3, \ldots$ there exists a finite covering of Ω by the balls $\left\{B\left(a_k^{(j)}; \frac{1}{k}\right)\right\}_{j=1}^{M_k}$. Construct $X^{(1)}$ by choosing points from Ω so that one point lies in each ball $B\left(a_k^{(j)}; 1\right)$. Construct $X^{(k+1)}$ by first choosing the points $X^{(k)}$ and then at least one extra point so that $X^{(k+1)}$ contains points from each ball $B\left(a_{k+1}^{(j)}; \frac{1}{k+1}\right)$.

Then $x \in \Omega \cap X^{(k)}$ implies $x \in B\left(a_k^{(j)}; \frac{1}{k}\right)$ for some j and hence $\operatorname{dist}\left(x, X^{(k)}\right) < \frac{1}{k}$. Hence $h_{X^{(k)}} = \sup_{x\in\Omega} \operatorname{dist}\left(x, X^{(k)}\right) < \frac{1}{k}$ and $\lim_{k\to\infty} h_{X^{(k)}} = 0$. ∎

The next theorem derives an upper bound for the uniform pointwise error of the minimal interpolant in terms of the product of the seminorm of the data function, a power of the sparsity of the independent data and various constants derived from the theory of Lagrange interpolation.

The order of uniform pointwise convergence is defined to be the power of the sparsity parameter, in this case $\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$ where θ and κ are the weight function parameters.

Theorem 152 *Let w be a weight function with properties W2 and W3 for order θ and smoothness parameter κ. Set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$.*

Suppose the notation and assumptions of Lemma 148 hold with the data point sparsity of X in Ω defined by $h_X = \sup_{\omega\in\Omega} \operatorname{dist}(\omega; X)$. Suppose also that $\mathcal{I}_X$ is the minimal interpolant on X of the data function $f_d \in X_w^{\theta}$.

Then there exist positive constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta}$ and

$$k_{\Omega,\theta,\eta} = \frac{d^{\eta/2}}{(2\pi)^{d/4}\sqrt{(2\eta)!}}\left(1+K'_{\Omega,\theta}\right) \text{ such that}$$

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\, k_{\Omega,\theta,\eta}\,(c_{\Omega,\theta}h_X)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq c_{\Omega,\theta}h_X}} \left|D^\beta G(y)\right|, \quad x \in \overline{\Omega}, \tag{4.71}$$

when $0 < h_X \leq h_{\Omega,\theta}$. *The Lagrange interpolation constants* $c_{\Omega,\theta}, K'_{\Omega,\theta}$ *and* $h_{\Omega,\theta}$ *are derived in Lemma 148.*

Further, $|f_d - \mathcal{I}_X f_d|_{w,\theta} \leq |f_d|_{w,\theta}$ *and the order of convergence is* η.

Finally we have the bound

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\, k_{\Omega,\theta,\eta}\,(\operatorname{diam}\Omega)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq \operatorname{diam}\Omega}} \left|D^\beta G(y)\right|, \quad x \in \overline{\Omega}.$$

Proof. Fix $x \in \Omega$. Since $\mathcal{I}_X$ interpolates the data from the function f_d it follows that $\mathcal{I}_X f(x) = f(x)$ when $x \in X$. Thus $\mathcal{P}(f_d - \mathcal{I}_X f_d) = 0$ and hence $f_d - \mathcal{I}_X f_d = \mathcal{Q}(f_d - \mathcal{I}_X f_d)$.

$$f_d(x) - \mathcal{I}_X f_d(x) = (f_d - \mathcal{I}_X f_d, R_x)_{w,\theta} = (\mathcal{Q}(f_d - \mathcal{I}_X f_d), \mathcal{Q}R_x)_{w,\theta} = \langle f_d - \mathcal{I}_X f_d, r_x\rangle_{w,\theta}.$$

But by part 4 of Theorem 125, $|r_x|_{w,\theta} = \sqrt{r_x(x)}$ and thus

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\,|r_x|_{w,\theta} = |f_d - \mathcal{I}_X f_d|_{w,\theta}\sqrt{r_x(x)}. \tag{4.72}$$

The upper bound for $\sqrt{r_x(x)}$ given by Theorem 150 is

$$\sqrt{r_x(x)} \leq \frac{d^{\eta/2}}{(2\pi)^{\frac{d}{4}}\sqrt{(2\eta)!}}\left(1+\sum_{k=1}^M |l_k(x)|\right)(\operatorname{diam} A_x)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq \operatorname{diam} A_x}} \left|D^\beta G(y)\right|,$$

so that by Lemma 148

$$\sqrt{r_x(x)} \leq \frac{d^{\eta/2}}{(2\pi)^{\frac{d}{4}}\sqrt{(2\eta)!}}\left(1+K'_{\Omega,\theta}\right)(c_{\Omega,\theta}h_X)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq c_{\Omega,\theta}h_X}} \left|(D^\beta G)(y)\right|, \tag{4.73}$$

when $h_X \leq h_{\Omega,\theta}$. Therefore

$$\begin{aligned}
|f_d(x) - \mathcal{I}_X f_d(x)| &\leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\frac{d^{\frac{\eta}{2}}\left(1+K'_{\Omega,\theta}\right)}{(2\pi)^{\frac{d}{4}}\sqrt{(2\eta)!}}(c_{\Omega,\theta}h_X)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq c_{\Omega,\theta}h_X,}} \left|(D^\beta G)(y)\right| \\
&\leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\, k_{\Omega,\theta,\eta}\,(c_{\Omega,\theta}h_X)^\eta \max_{\substack{|\beta|=2\eta \\ |y|\leq c_{\Omega,\theta}h_X}} \left|(D^\beta G)(y)\right|, \quad x \in \Omega,
\end{aligned}$$

when $h_X \leq h_{\Omega,\theta}$. The last inequality can be extended to $\overline{\Omega}$ because the inequality is uniform on Ω, Ω is bounded and $f_d - \mathcal{I}_X f_d$ is continuous on $\mathbb{R}^d$. That $|f_d - \mathcal{I}_X f_d|_{w,\theta} \leq |f_d|_{w,\theta}$ was shown in part 2 of Theorem 147.

The final bound is true since $\operatorname{diam} A_x \leq \operatorname{diam}\Omega$. ∎

Remark 153 *If* $\kappa < 1/2$ *then* $\eta = 0$ *and the last theorem only shows the interpolation error is bounded i.e. zero order of convergence. In the next section we show how to overcome this limitation and obtain positive convergence orders for these cases and improved orders of convergence when* $\kappa \geq 1/2$. *In the process we will demonstrate improved order of convergence estimates for almost all thin-plate splines and shifted thin-plate splines.*

4.11 Slightly improved convergence results

Noting Remark 153 to the last lemma in the previous section, we will show how to obtain improved interpolant convergence estimates using a Taylor series expansion result for distributions. Our results will be illustrated using the thin-plate splines and the shifted thin-plate splines. The improvement in convergence order for these examples is at most 1/2.

To prove Lemma 155 we will require the following result:

Lemma 154 *Suppose $u \in L^1_{loc}\left(\mathbb{R}^d\right)$ and $\phi \in C_0^\infty$. Then for all $b \in \mathbb{R}^d$,*

$$\int_0^1 \int |u(x+tb)| \quad |\phi(x)|\, dxdt < \infty. \tag{4.74}$$

Proof. Suppose $\operatorname{supp}\phi \subset B(0;R)$. Then

$$\begin{aligned}
\int_0^1 \int |u(x+tb)| \quad |\phi(x)|\, dxdt = \int_0^1 \int |u(y)| \quad |\phi(y-tb)|\, dydt &= \int_0^1 \int_{|y-tb|\le R} |u(y)| \quad |\phi(y-tb)|\, dydt \\
&\le \int_0^1 \int_{|y|\le R+|b|} |u(y)| \quad |\phi(y-tb)|\, dydt \\
&\le \|\phi\|_\infty \int_0^1 \int_{|y|\le R+|b|} |u(y)|\, dydt \\
&\le \|\phi\|_\infty \int_{|y|\le R+|b|} |u(y)|\, dy \\
&< \infty.
\end{aligned}$$

■

Our Taylor series distribution result is:

Lemma 155 *Suppose $u \in C^{(n-1)}\left(\mathbb{R}^d\right)$ and the distributional derivatives $\left\{D^\beta u\right\}_{|\beta|=n}$ are L^1_{loc} functions. Suppose also that for each fixed $b \neq 0$ the integrals*

$$\int_0^1 (1-t)^{n-1} \left|(D^\beta u)(z+tb)\right| dt, \quad |\beta| = n;\ z, b \in \mathbb{R}^d, \tag{4.75}$$

have polynomial growth in z. Then

$$u(z+b) = \sum_{|\beta|<n} \frac{b^\beta}{\beta!} \left(D^\beta u\right)(z) + \left(\mathcal{R}_n u\right)(z,b), \quad z, b \in \mathbb{R}^d, \tag{4.76}$$

where $\mathcal{R}_n u$ is the integral remainder term

$$\left(\mathcal{R}_n u\right)(z,b) = n \sum_{|\beta|=n} \frac{b^\beta}{\beta!} \int_0^1 (1-t)^{n-1} (D^\beta u)(z+tb)dt, \quad z, b \in \mathbb{R}^d. \tag{4.77}$$

Proof. In order to overcome the fact that $D^\beta u$ may not be $C^{(0)}\left(\mathbb{R}^d\right)$ when $|\beta| = n$, we will use a Taylor series expansion with remainder for distributions. Suppose $\phi \in C_0^\infty$. Then the conditions on u allow us to use Lemma 154 to show that the iterated integrals 4.79 and 4.80 are absolutely convergent and thus apply Fubini's theorem to swap the orders of integration in the following calculations:

$$\begin{aligned}
\left[u(z+b), \phi(z)\right] &= \left[u(z), \phi(z-b)\right] \\
&= \left[u(z), \sum_{|\beta|\le n} \frac{(-b)^\beta}{\beta!} D^\beta \phi(z)\right] + \\
&\quad + \left[u(z), n \sum_{|\beta|=n} \frac{(-b)^\beta}{\beta!} \int_0^1 (1-t)^{n-1} \left(D^\beta \phi\right)(z+tb)\, dt\right] \\
&= \sum_{|\beta|<n} \frac{b^\beta}{\beta!} \left[D^\beta u, \phi\right] + \\
&\quad + n \sum_{|\beta|=n} \frac{(-b)^\beta}{\beta!} \left[u, \int_0^1 (1-t)^{n-1} \left(D^\beta \phi\right)(\cdot - tb)\, dt\right].
\end{aligned} \tag{4.78}$$

We now analyze the integral remainder term of 4.78:

$$\begin{aligned}
\left[u, \int_0^1 (1-t)^{n-1} \left(D^\beta \phi\right)(\cdot - tb)\, dt\right] &= \int u(z) \int_0^1 (1-t)^{n-1} \left(D^\beta \phi\right)(z - tb)\, dt\, dz \\
&= \int_0^1 (1-t)^{n-1} \int u(z) \left(D^\beta \phi\right)(z - tb)\, dz\, dt \qquad (4.79) \\
&= \int_0^1 (1-t)^{n-1} \left[u(z), \left(D^\beta \phi\right)(z - tb)\right]\, dt \\
&= \int_0^1 (1-t)^{n-1} \left[u(z + tb), \left(D^\beta \phi\right)(z)\right]\, dt \\
&= (-1)^{|\beta|} \int_0^1 (1-t)^{n-1} \left[\left(D^\beta u\right)(z + tb), \phi(z)\right]\, dt \\
&= (-1)^{|\beta|} \int_0^1 (1-t)^{n-1} \int \left(D^\beta u\right)(z + tb)\, \phi(z)\, dz\, dt \\
&= (-1)^{|\beta|} \int \int_0^1 (1-t)^{n-1} \left(D^\beta u\right)(z + tb)\, dt\, \phi(z)\, dz \qquad (4.80) \\
&= (-1)^{|\beta|} \left[\int_0^1 (1-t)^{n-1} \left(D^\beta u\right)(z + tb)\, dt, \phi(z)\right],
\end{aligned}$$

so 4.78 now becomes

$$\begin{aligned}
\left[u(z+b), \phi(z)\right] &= \sum_{|\beta|<n} \frac{b^\beta}{\beta!} \left[D^\beta u(z), \phi(z)\right] + \\
&\quad + n \sum_{|\beta|=n} \frac{b^\beta}{\beta!} \left[\int_0^1 (1-t)^{n-1} \left(D^\beta u\right)(z + tb)\, dt, \phi(z)\right],
\end{aligned}$$

for $\phi \in C_0^\infty$. Thus

$$u(z+b) = \sum_{|\beta|<n} \frac{b^\beta}{\beta!} \left(D^\beta u\right)(z) + n \sum_{|\beta|=n} \frac{b^\beta}{\beta!} \int_0^1 (1-t)^{n-1} \left(D^\beta u\right)(z + tb)\, dt,$$

as claimed. ■

Using the previous lemma we now modify Theorem 150 to obtain the following estimate for $\sqrt{r_x(x)}$ which requires the assumptions 4.81 and 4.82 to be valid.

Theorem 156 *Suppose w is a weight function with properties W2 and W3 for order θ and κ. Set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Also suppose G is a basis function of order θ such that the distributions $\left\{D^\beta G\right\}_{|\beta|=2\eta+1}$ are L^1_{loc} functions such that for each fixed $b \neq 0$ the integrals*

$$\int_0^1 (1-t)^{2\eta} \left|(D^\beta G)(z + tb)\right| dt, \quad z, b \in \mathbb{R}^d, \ |\beta| = 2\eta + 1, \qquad (4.81)$$

have polynomial growth in z.

Further, suppose there exist constants $r_G, c_{G,\eta} > 0$ and $\delta_G \geq 0$ such that

$$\left|b^\beta\right| \int_0^1 (1-t)^{2\eta} \left|(D^\beta G)(tb)\right| dt \leq \frac{c_{G,\eta}}{2\sigma} |b|^{2(\eta+\delta_G)}, \quad |b| \leq r_G, \ |\beta| = 2\eta + 1, \qquad (4.82)$$

where $\sigma = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa + 1 \rfloor\right\}$.

Regarding unisolvency, assume $A = \{a_k\}_{k=1}^M$ is a minimal θ-unisolvent set and that $\{l_k\}_{k=1}^M$ is the corresponding unique cardinal basis for P_θ. Now construct $\mathcal{P}, \mathcal{Q}, R_x$ using A and $\{l_k\}_{k=1}^M$.

Then if $r_x = \mathcal{Q} R_x$ we have the estimate

$$\sqrt{r_x(x)} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + \sum_{k=1}^M |l_k(x)|\right) (\operatorname{diam} A_x)^{\eta+\delta_G}, \quad \operatorname{diam} A_x \leq r_G, \ x \in \Omega, \qquad (4.83)$$

where $A_x = A \cup \{x\}$ and $c_{G,\eta,\sigma} = \frac{d^{\lceil \sigma \rceil}}{(2\pi)^{\frac{d}{2}} \lceil \sigma \rceil!} c_{G,\eta}$.

Proof. The proof will follow that of Theorem 150. From part 4 of Theorem 125, $r_x(y) = \langle r_x, r_y \rangle_{w,\theta}$ so that $r_x(x) = |r_x|^2_{w,\theta} \geq 0$. Also from Theorem 125, $r_x(y) = (2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x G(y-x)$ when $x \neq y$.

Since $G \in C_{BP}^{(\lfloor 2\kappa \rfloor)}$ and the functions $\left\{D^\beta G\right\}_{|\beta| = \lfloor 2\eta+1 \rfloor}$ are L^1_{loc} conditions 4.81 mean we can use Lemma 155 to expand G about the origin and get

$$G(y-x) = \sum_{|\beta|<2\sigma} \frac{(y-x)^\beta}{\beta!} \left(D^\beta G\right)(0) + \mathcal{R}_{2\sigma}(0, y-x), \tag{4.84}$$

where $2\sigma = \min\{2\theta, \lfloor 2\kappa \rfloor + 1\}$.

To calculate $r_x(y)$ apply the operator $\mathcal{Q}_y \mathcal{Q}_x$ to 4.84 and then noting that $2\sigma \leq 2\theta$ use Theorem 118 to eliminate the power series and obtain

$r_x(y) = (2\pi)^{-\frac{d}{2}} \mathcal{Q}_y \mathcal{Q}_x \mathcal{R}_{2\sigma}(0, y-x)$. Expanding $\mathcal{Q}_y$ and $\mathcal{Q}_x$ using $\mathcal{P}_y$ and $\mathcal{P}_x$ now gives

$$\begin{aligned}
(2\pi)^{\frac{d}{2}} r_x(y) &= \mathcal{R}_{2\sigma}(0, y-x) - \mathcal{P}_x\left(\mathcal{R}_{2\sigma}(0, y-x)\right) - \mathcal{P}_y\left(\mathcal{R}_{2\sigma}(0, y-x)\right) + \mathcal{P}_y \mathcal{P}_x\left(\mathcal{R}_{2\sigma}(0, y-x)\right) \\
&= \mathcal{R}_{2\sigma}(0, y-x) - \sum_{j=1}^M \mathcal{R}_{2\sigma}(0, y-a_j)\, l_j(x) - \sum_{k=1}^M \mathcal{R}_{2\sigma}(0, a_k - x)\, l_k(y) + \\
&\quad + \sum_{j,k=1}^M \mathcal{R}_{2\sigma}(0, a_k - a_j)\, l_j(x)\, l_k(y),
\end{aligned}$$

so that

$$\begin{aligned}
(2\pi)^{\frac{d}{2}} r_x(x) &\leq \sum_{j=1}^M |\mathcal{R}_{2\sigma}(0, x-a_j)|\, |l_j(x)| + \sum_{k=1}^M |\mathcal{R}_{2\sigma}(0, a_k - x)|\, |l_k(x)| + \\
&\quad + \sum_{j,k=1}^M |\mathcal{R}_{2\sigma}(0, a_k - a_j)|\, |l_j(x)|\, |l_k(x)| \\
&\leq 2\left(\sum_{k=1}^M |l_k(x)|\right) \max_k |\mathcal{R}_{2\sigma}(0, x-a_k)| + \left(\sum_{k=1}^M |l_k(x)|\right)^2 \max_{j,k} |\mathcal{R}_{2\sigma}(0, a_k - a_j)|.
\end{aligned}$$

Set $\tau = \eta + \delta_G$. Then from Lemma 155

$$\mathcal{R}_{2\sigma}(0, b) = 2\sigma \sum_{|\beta|=2\sigma} \frac{b^\beta}{\beta!} \int_0^1 (1-t)^{2\sigma-1} (D^\beta G)(tb)dt,$$

and so when $|b| \leq r_G$ estimate 4.82 implies

$$\begin{aligned}
|\mathcal{R}_{2\sigma}(0, b)| \leq \left(\sum_{|\beta|=2\sigma} \frac{1}{\beta!}\right) c_{G,\eta} |b|^{2\tau} &= \left(\sum_{|\beta|=2\sigma} \frac{\mathbf{1}^{2\beta}}{\beta!}\right) c_{G,\eta} |b|^{2\tau} \leq \left(\sum_{|\beta|=2\lceil\sigma\rceil} \frac{\mathbf{1}^{2\beta}}{\beta!}\right) c_{G,\eta} |b|^{2\tau} \\
&= \frac{|\mathbf{1}|^{2\lceil\sigma\rceil}}{\lceil\sigma\rceil!} c_{G,\eta} |b|^{2\tau} = \frac{d^{\lceil\sigma\rceil}}{\lceil\sigma\rceil!} c_{G,\eta} |b|^{2\tau}.
\end{aligned}$$

Hence if $\operatorname{diam} A_x \leq r_G$

$$\begin{aligned}
(2\pi)^{\frac{d}{2}} r_x(x) &\leq \frac{d^{\lceil\sigma\rceil}}{\lceil\sigma\rceil!} c_{G,\eta} 2\left(\sum_{k=1}^M |l_k(x)|\right)\left(\max_k |x-a_k|\right)^{2\tau} + \\
&\quad + \frac{d^{\lceil\sigma\rceil}}{\lceil\sigma\rceil!} c_{G,\eta} \left(\sum_{k=1}^M |l_k(x)|\right)^2 \left(\max_{j,k} |a_j - a_k|\right)^{2\tau} \\
&< \frac{d^{\lceil\sigma\rceil}}{\lceil\sigma\rceil!} c_{G,\eta} \left(1 + \sum_{k=1}^M |l_k(x)|\right)^2 (\operatorname{diam} A_x)^{2\tau},
\end{aligned}$$

so that

$$\begin{aligned}\sqrt{r_x(x)} &\leq \frac{d^{\frac{\lceil\sigma\rceil}{2}}\sqrt{c_{G,\eta}}}{(2\pi)^{\frac{d}{4}}\sqrt{\lceil\sigma\rceil!}}\left(1+\sum_{k=1}^{M}|l_k(x)|\right)(\operatorname{diam}A_x)^{\eta+\delta_G} \\ &= \sqrt{c_{G,\eta,\sigma}}\left(1+\sum_{k=1}^{M}|l_k(x)|\right)(\operatorname{diam}A_x)^{\eta+\delta_G},\end{aligned}$$

as claimed. ■

We can now prove our improved interpolant convergence estimate.

Corollary 157 *Suppose the notation and assumptions of Lemma 148 and Theorem 156 hold. Suppose also that $\mathcal{I}_X$ is the minimal interpolant on X of the data function $f_d \in X_w^\theta$.*

Then there exist positive constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta}$ such that

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\sqrt{c_{G,\eta,\sigma}}\left(1+K'_{\Omega,\theta}\right)(c_{\Omega,\theta}h_X)^{\eta+\delta_G}, \quad x \in \overline{\Omega}, \tag{4.85}$$

when $0 < h_X \leq \min\{h_{\Omega,\theta}, r_\Omega\}$. The constants $c_{\Omega,\theta}, K'_{\Omega,\theta}$ and $h_{\Omega,\theta}$ are derived in Lemma 148.

Next, $|f_d - \mathcal{I}_X f_d|_{w,\theta} \leq |f_d|_{w,\theta}$ so the order of convergence is $\eta + \delta_G$.

Finally, for all independent data X

$$|f_d(x) - \mathcal{I}_X f_d(x)| \leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\, k_{\Omega,\theta,\eta}\left(\min\{\operatorname{diam}\Omega, r_G\}\right)^{\eta+\delta_G}, \quad x \in \overline{\Omega}. \tag{4.86}$$

Proof. If $x \in \Omega$ and $\operatorname{diam} A_x \leq r_G$ then from 4.72, 4.83 and Lemma 148

$$\begin{aligned}|f_d(x) - \mathcal{I}_X f_d(x)| &\leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\sqrt{r_x(x)} \\ &\leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\frac{d^{\frac{\lceil\sigma\rceil}{2}}\sqrt{c_{G,\eta}}}{(2\pi)^{\frac{d}{4}}\sqrt{\lceil\sigma\rceil!}}\left(1+\sum_{k=1}^{M}|l_k(x)|\right)(\operatorname{diam}A_x)^{\eta+\delta_G} \\ &\leq |f_d - \mathcal{I}_X f_d|_{w,\theta}\frac{d^{\frac{\lceil\sigma\rceil}{2}}\sqrt{c_{G,\eta}}}{(2\pi)^{\frac{d}{4}}\sqrt{\lceil\sigma\rceil!}}\left(1+K'_{\Omega,\theta}\right)(\operatorname{diam}A_x)^{\eta+\delta_G} \\ &= |f_d - \mathcal{I}_X f_d|_{w,\theta}\sqrt{c_{G,\eta,\sigma}}\left(1+K'_{\Omega,\theta}\right)(\operatorname{diam}A_x)^{\eta+\delta_G}.\end{aligned}$$

Now by Lemma 148 if $h_X \leq h_{\Omega,\theta}$ then there exists a minimal unisolvent $A \subset \Omega$ such that $\operatorname{diam} A_x \leq c_{\Omega,\theta}h_X$ which proves 4.85. Finally, the estimate 4.86 is true since $\operatorname{diam} A_x \leq \operatorname{diam}\Omega$. ■

Lemma 158 *Suppose for a given order θ the weight function w has property W3 for some κ. Then one of the following must hold:*

1. *For some $s > \kappa$, w has property W3 for all $0 \leq \kappa < s$ and does not have property W3 for $\kappa \geq s$;*
2. *For some $s \geq \kappa$, w has property W3 for all $0 \leq \kappa \leq s$ and does not have property W3 for $\kappa > s$;*
3. *w has property W3 for all $\kappa \geq 0$ i.e. $s = \infty$.*

Proof. This result follows from the fact that if property W3 holds for $\kappa = \kappa_0$ then W3 holds for all $0 \leq \kappa \leq \kappa_0$. ■

The examples of this section will have either property 1 or property 3 of the last lemma and these examples will be concerned with finding the 'maximum' order of convergence given the constraint applied to κ. Now $\eta = \max\limits_{0\leq\kappa<s}\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa\rfloor\right\}$ is the maximum order of convergence defined by Theorem 150. To this is added an increment of convergence order δ_G defined by Theorem 156. The following result will be used to calculate η and σ when the weight function has property 1 or 3 of the last lemma.

Theorem 159 *Given $s > 0$ suppose w is a weight function with properties W2 and W3 for order θ and all $\kappa < s$. Set $\eta = \max\limits_{0\leq\kappa<s}\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa\rfloor\right\}$ and $\sigma = \max\limits_{0\leq\kappa<s}\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa+1\rfloor\right\}$.*

Then

$$\eta = \begin{cases}\min\left\{\theta, \frac{1}{2}\lfloor 2s\rfloor\right\}, & if\ 2s\ is\ not\ an\ integer, \\ \min\left\{\theta, s-\frac{1}{2}\right\}, & if\ 2s\ is\ an\ integer.\end{cases} \tag{4.87}$$

If $s \leq \theta$ then

$$\eta = \begin{cases}\frac{1}{2}\lfloor 2s\rfloor, & if\ 2s\ is\ not\ an\ integer, \\ s-\frac{1}{2}, & if\ 2s\ is\ an\ integer.\end{cases} \tag{4.88}$$

If $s > \theta$ then

$$\eta = \theta. \tag{4.89}$$

Regarding σ:

$$\sigma = \begin{cases} \min\left\{\theta, \frac{1}{2}\lfloor 2s+1 \rfloor\right\}, & \textit{if } 2s \textit{ is not an integer}, \\ \min\{\theta, s\}, & \textit{if } 2s \textit{ is an integer}, \end{cases} \tag{4.90}$$

and σ is related to η by

$$\sigma = \begin{cases} \eta, & \lfloor 2s \rfloor \geq 2\theta, \\ \eta + \frac{1}{2}, & \lfloor 2s \rfloor \leq 2\theta - 1. \end{cases} \tag{4.91}$$

Proof. Let $\eta_\kappa = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Suppose $2s$ is not an integer and $0 < \varepsilon < 2s - \lfloor 2s \rfloor$. Then

$$\begin{aligned} \eta = \max\left\{\eta_\kappa : \left(1 - \frac{\varepsilon}{2s}\right) s \leq \kappa < s\right\} \geq \min\left\{\theta, \frac{1}{2}\left\lfloor 2\left(1 - \frac{\varepsilon}{2s}\right) s\right\rfloor\right\} &= \min\left\{\theta, \frac{1}{2}\lfloor 2s - \varepsilon \rfloor\right\} \\ &\geq \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor\right\}, \end{aligned}$$

but $\eta = \max\{\eta_\kappa : 0 \leq \kappa < s\} \leq \max\{\eta_\kappa : 0 \leq \kappa \leq s\} \leq \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor\right\}$ so when $2s$ is not an integer it follows that $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor\right\}$.

If $2s$ is an integer then

$$\begin{aligned} \eta = \max\{\eta_\kappa : 0 \leq \kappa < s\} &= \max\{\eta_\kappa : 0 \leq 2\kappa < 2s\} \\ &= \max\{\eta_\kappa : 2s - 1 < 2\kappa < 2s\} \\ &= \max\left\{\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\} : 2s - 1 < 2\kappa < 2s\right\} \\ &= \max\left\{\min\left\{\theta, \frac{1}{2}(2s-1)\right\} : 2s - 1 < 2\kappa < 2s\right\} \\ &= \min\left\{\theta, s - \frac{1}{2}\right\}. \end{aligned}$$

To prove 4.88 suppose $s \leq \theta$: If $2s$ is not an integer then by 4.87, $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor\right\}$. But $\frac{1}{2}\lfloor 2s \rfloor \leq \frac{1}{2}\lfloor 2\theta \rfloor = \theta$ so $\eta = \frac{1}{2}\lfloor 2s \rfloor$. If $2s$ is an integer then again by 4.87, $\eta = \min\left\{\theta, s - \frac{1}{2}\right\} = s - \frac{1}{2}$.

To prove 4.89 suppose $s > \theta$: If $2s$ is not an integer then $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor\right\} = \theta$. If $2s$ is an integer then $\eta = \min\left\{\theta, s - \frac{1}{2}\right\}$. But $2s > 2\theta$ implies $2s - 1 \geq 2\theta$ i.e. $s - \frac{1}{2} \geq \theta$ and $\eta = \theta$.

Let $\sigma_\kappa = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa + 1 \rfloor\right\}$. Suppose $2s$ is not an integer and $0 < \varepsilon < 2s - \lfloor 2s \rfloor$. Then

$$\begin{aligned} \sigma = \max\left\{\sigma_\kappa : \left(1 - \frac{\varepsilon}{2s}\right) s \leq \kappa < s\right\} &\geq \min\left\{\theta, \frac{1}{2}\left\lfloor 2\left(1 - \frac{\varepsilon}{2s}\right) s\right\rfloor + \frac{1}{2}\right\} \\ &= \min\left\{\theta, \frac{1}{2}\lfloor 2s - \varepsilon \rfloor + \frac{1}{2}\right\} \\ &\geq \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor + \frac{1}{2}\right\}, \end{aligned}$$

but $\sigma = \max\{\sigma_\kappa : 0 \leq \kappa < s\} \leq \max\{\sigma_\kappa : 0 \leq \kappa \leq s\} \leq \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor + \frac{1}{2}\right\}$ so when $2s$ is not an integer it follows that $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2s \rfloor + \frac{1}{2}\right\} = \min\left\{\theta, \frac{1}{2}\lfloor 2s + 1 \rfloor\right\}$.

If $2s$ is an integer then

$$\begin{aligned} \sigma = \max\{\sigma_\kappa : 0 \leq \kappa < s\} &= \max\{\sigma_\kappa : 0 \leq 2\kappa < 2s\} \\ &= \max\{\sigma_\kappa : 2s - 1 < 2\kappa < 2s\} \\ &= \max\left\{\min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor + \frac{1}{2}\right\} : 2s - 1 < 2\kappa < 2s\right\} \\ &= \max\left\{\min\left\{\theta, \frac{1}{2}(2s-1) + \frac{1}{2}\right\} : 2s - 1 < 2\kappa < 2s\right\} \\ &= \min\{\theta, s\}, \end{aligned}$$

which proves 4.90.

Finally, 4.91 is proved by using 4.90 and 4.87 to calculate σ and η for $\lfloor 2s\rfloor \geq 2\theta$ and $\lfloor 2s\rfloor \leq 2\theta-1$. ■

Next we will illustrate the theory of this subsection by using as examples the radial thin-plate splines and shifted thin-plate splines which were among the positive order basis functions studied in Chapter 2.

Example 1: The thin-plate splines

We will derive the combinations of η and δ_G for which the thin-plate spline basis functions satisfy the requirements of Theorem 156. Corollary 157 then tells us the orders of convergence of the interpolant are $\eta+\delta_G$, which is an improvement of δ_G.

The thin-plate spline weight and basis functions were studied in Subsection 2.3.1 of Chapter 2. By Theorem 71 the equation

$$w(\xi)=\frac{1}{e(s)}\left|\xi\right|^{-2\theta+2s+d}, \tag{4.92}$$

defines a thin-plate spline weight function with properties W2.1 and W3.2 with positive integer order θ and non-negative κ iff $\kappa<s<\theta$. The corresponding basis functions are defined by

$$G(x)=\begin{cases}(-1)^{s+1}\left|x\right|^{2s}\log\left|x\right|, & s=1,2,3,\ldots,\\ (-1)^{\lceil s\rceil}\left|x\right|^{2s}, & s>0,\ s\neq 1,2,3,\ldots.\end{cases} \tag{4.93}$$

Now suppose $\kappa<s<\theta$.

$\boxed{\textbf{Case 1 } 0<s<\theta,\ s\neq 1,2,3,\ldots}$ By 4.93, $G(x)=(-1)^{\lceil s\rceil}\left|x\right|^{2s}$. This is a homogeneous function of degree $2s$ and hence for some constants $k_{s,n}=\max\limits_{|\beta|=n}\max\limits_{|x|=1}\left|D^{\beta}G(x)\right|$,

$$\left|D^{\beta}G(x)\right|\leq k_{s,|\beta|}\left|x\right|^{2s-|\beta|},\quad \beta\geq 0, \tag{4.94}$$

and so $D^{\beta}G\in L^1_{loc}$ iff

$$|\beta|-2s<d, \tag{4.95}$$

and for all $\beta\geq 0$

$$\int_0^1(1-t)^{2\eta}\left|(D^{\beta}G)(z+tb)\right|dt\leq k_{s,|\beta|}\int_0^1\left|z+tb\right|^{2s-|\beta|}dt. \tag{4.96}$$

We want to apply Theorem 156 which assumes $|\beta|=2\eta+1$. From Theorem 159, $\eta=\frac{1}{2}\lfloor 2s\rfloor$ so $2s-|\beta|=2s-2\eta-1=2s-\lfloor 2s\rfloor-1$ and $-1<2s-|\beta|\leq 0$ which means 4.95 is satisfied. Hence

$$\begin{aligned}\int_0^1\left|z+bt\right|^{2s-|\beta|}dt\leq\int_0^1\left|\left|z\right|-\left|b\right|t\right|^{2s-|\beta|}dt&=\left|b\right|^{2s-|\beta|}\int_0^1\left|\frac{|z|}{|b|}-t\right|^{2s-|\beta|}dt\\&=\left|b\right|^{2s-|\beta|}\int_{|z||b|^{-1}-1}^{|z||b|^{-1}}\left|u\right|^{2s-|\beta|}du\\&<2\left|b\right|^{2s-|\beta|}\int_0^{1+|z||b|^{-1}}u^{2s-|\beta|}du\\&\leq\frac{2\left|b\right|^{2s-|\beta|}}{2s-|\beta|+1}\left(1+\frac{|z|}{|b|}\right)^{2s-|\beta|+1}\\&<\infty,\end{aligned}$$

and 4.96 implies

$$\int_0^1(1-t)^{2\eta}\left|(D^{\beta}G)(z+tb)\right|dt\leq\frac{2k_{s,|\beta|}}{2s-|\beta|+1}\left(1+\frac{|z|}{|b|}\right)^{2s-|\beta|+1},$$

which in turn implies polynomial growth in $|z|$.

When $z=0$, 4.96 implies

$$\begin{aligned}|b|^{|\beta|}\int_0^1 (1-t)^{2\eta}\left|(D^\beta G)(tb)\right|dt = k_{s,|\beta|}\,|b|^{|\beta|}\int_0^1 |tb|^{2s-|\beta|}\,dt &= \frac{1}{2}k_{s,|\beta|}\,|b|^{2s}\int_0^1 t^{2s-|\beta|}dt \\ &= \frac{k_{s,|\beta|}}{2\,(2s+1-|\beta|)}\,|b|^{2s} \\ &= \frac{k_{s,|\beta|}}{2\,(2s+1-|\beta|)}\,|b|^{2s}.\end{aligned}$$

We want an estimate of the form 4.82. But from Theorem 159, $\eta=\frac{1}{2}\lfloor 2s\rfloor$ so $2s=2\eta+(2s-\lfloor 2s\rfloor)$ and we can write $2s=2\eta+2\delta_G$ where $\delta_G=s-\frac{1}{2}\lfloor 2s\rfloor<\frac{1}{2}$. Hence valid values of δ_G are $0\le\delta_G\le s-\frac{1}{2}\lfloor 2s\rfloor$. Note that $0\le s-\frac{1}{2}\lfloor 2s\rfloor<\frac{1}{2}$ and that $s-\frac{1}{2}\lfloor 2s\rfloor=0$ iff $s=n$ or $n+\frac{1}{2}$ for some integer $n\ge 0$.

$\boxed{\textbf{Case 2}\ s=1,2,3,\ldots,\theta-1}$ To apply Theorem 156 we need to estimate the derivatives $\left\{D^\beta G\right\}_{|\beta|=2\eta+1}$. But from Theorem 159, $\eta=s-\frac{1}{2}$ so that $|\beta|=2\eta+1=2s$.

Next, 4.93 implies $G(x)=(-1)^{s+1}|x|^{2s}\log|x|$ for $s=1,2,3,\ldots,\theta-1$. But any first order derivative of $\log|x|$ is a homogeneous function of order -1 so there exist constants $k_{s,|\beta|}$ and $k'_{s,|\beta|}$ such that

$$\left|D^\beta G(x)\right|\le k_{s,|\beta|}\,|x|^{2s-|\beta|}\,|\log|x||+k'_{s,|\beta|}\,|x|^{2s-|\beta|},\quad 0\le|\beta|\le 2s,$$

and when $|\beta|=2s$

$$\left|D^\beta G(x)\right|\le k_{s,2s}\,|\log|x||+k'_{s,2s},\quad |\beta|=2s. \tag{4.97}$$

Now $|\log|x||\in L^1_{loc}$ iff $\int_{|\cdot|\le 1}|\log|x||\,dx<\infty$ and by using spherical polar coordinates (r,ϕ) e.g. Section 5.43 Adams [2], it is straight forward to show that this integral exists because $r^d\,|\log r|\in L^1_{loc}\left(\mathbb{R}^1\right)$.

Further, the integrals 4.81 satisfy the inequalities

$$\begin{aligned}\int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(z+tb)\right|dt &\le \int_0^1\left|\left(D^\beta G\right)(z+tb)\right|dt \\ &\le k_{s,2s}\int_0^1|\log|z+tb||\,dt+k'_{s,2s}\,|b|^{2s}\int_0^1 dt \\ &= k_{s,2s}\int_0^1|\log|z+tb||\,dt+k'_{s,2s}\,|b|^{2s}.\end{aligned} \tag{4.98}$$

Suppose $d=1$. Then if $|z+b|\le 1$ and $|z|\le 1$ then $\int_0^1|\log|z+tb||\,dt=\frac{1}{b}\int_z^{z+b}|\log|\tau||\,d\tau\le\frac{1}{b}\int_{-1}^1|\log|\tau||\,d\tau=\frac{2}{b}$. Otherwise, if $|z+b|\ge 1$ or $|z|\ge 1$ then set $a_1=\max\{|z+b|,|z|\}$ so that

$$\begin{aligned}\int_0^1|\log|z+tb||\,dt=\frac{1}{b}\int_z^{z+b}|\log|\tau||\,d\tau\le\frac{1}{b}\int_{-1}^1|\log|\tau||\,d\tau+\frac{2}{b}\int_1^{a_1}|\log|\tau||\,d\tau &\le\frac{2}{b}+\frac{2}{b}\int_1^{a_1}a_1d\tau \\ &\le\frac{2}{b}+\frac{2}{b}|a_1|^2 \\ &\le\frac{2}{b}+\frac{2}{b}|z+b|^2,\end{aligned}$$

proving that when $d=1$ the integrals 4.81 have polynomial increase in z for any given $b\ne 0$.

Now assume $d>1$. To evaluate $\int_0^1|\log|z+tb||\,dt$ first observe that the interval $[z,z+b]$ can be rotated so that $b_1>0$ and $b_i=0$ for $i\ge 2$. To this end set $z=(z_1,z')$ so that

$$\begin{aligned}\int_0^1|\log|z+tb||\,dt=\int_0^1|\log|(z_1+b_1t,z')||\,dt &= \frac{1}{2}\int_0^1\left|\log\left((z_1+b_1t)^2+|z'|^2\right)\right|dt. \\ &= \frac{1}{2}\int_0^1\left|\log\left((z_1+b_1t)^2+|z'|^2\right)\right|dt.\end{aligned} \tag{4.99}$$

If $|z'|\ge 1$ then z and $z+b$ both lie outside the unit sphere and

$$\int_0^1|\log|z+tb||\,dt\le\frac{1}{2}\int_0^1\left((z_1+b_1)^2+z_2^2\right)dt=\frac{1}{2}\left((z_1+b_1)^2+z_2^2\right)\le|z|^2+|b|^2. \tag{4.100}$$

If $|z'| < 1$ and z and $z + b$ both lie inside the unit sphere we have $(z_1 + b_1 t)^2 + |z'|^2 < 1$ and

$$\begin{aligned}
\int_0^1 |\log|z + tb|| \, dt \leq -\frac{1}{2b_1} \int_0^1 \log\left(|z_1 + b_1 t|^2\right) dt &= -\frac{1}{|b|} \int_{z_1}^{z_1+b_1} \log|s| \, ds \\
&\leq -\frac{2}{|b|} \int_0^1 \log s ds \\
&= \frac{2}{|b|}. \qquad (4.101)
\end{aligned}$$

If $|z'| < 1$ and z and $z + b$ do not both lie inside the unit sphere set $a_1 = \max\{|z_1|, |z_1 + b_1|\} \geq 1$ so that by 4.99

$$\begin{aligned}
\int_0^1 |\log|z + tb|| \, dt &\leq \frac{1}{2} \int_0^1 \left|\log\left((z_1 + b_1 t)^2 + |z'|^2\right)\right| dt \\
&\leq \frac{1}{2b_1} \int_{z_1}^{z_1+b_1} \left|\log\left(\tau^2 + |z'|^2\right)\right| d\tau \\
&= -\frac{1}{|b|} \int_0^{\sqrt{1-z_2^2}} \log\left(\tau^2 + |z'|^2\right) d\tau + \frac{1}{|b|} \int_{\sqrt{1-z_2^2}}^{a_1} \log\left(\tau^2 + |z'|^2\right) d\tau \\
&\leq -\frac{2}{|b|} \int_0^1 \log \tau d\tau + \frac{1}{|b|} \int_{\sqrt{1-z_2^2}}^{a_1} \left(a_1^2 + |z'|^2\right) d\tau \\
&\leq -\frac{2}{|b|} \int_0^1 \log \tau d\tau + \frac{1}{|b|} \int_0^{a_1} \left(a_1^2 + |z'|^2\right) d\tau \\
&= \frac{2}{|b|} + \frac{1}{|b|} a_1 \left(a_1^2 + |z'|^2\right),
\end{aligned}$$

and since $\int_0^1 \log \tau d\tau = 1$, $a_1 \leq |z| + |b|$ and $|z'|^2 \leq |z|^2$ it follows that $\int_0^1 |\log|z + tb|| \, dt$ has polynomial increase in z. Indeed, combining this result with the estimates 4.101 and 4.100 we can conclude that the integrals 4.81 have polynomial increase in z when $d > 1$.

Finally, when $z = 0$ choose $r_G = 1$. Then 4.98 implies that when $|b| \leq r_G$ and $0 < \varepsilon \leq s$:

$$\begin{aligned}
|b|^{|\beta|} \int_0^1 (1-t)^{2\eta} \left|\left(D^\beta G\right)(tb)\right| dt &\leq k_{s,2s} |b|^{2s} \int_0^1 |\log|tb|| \, dt + k'_{s,2s} |b|^{2s} \\
&= k_{s,2s} |b|^{2s} \left(\int_0^1 \log\frac{1}{t} dt + \int_0^1 |\log|b|| \, dt\right) + k'_{s,2s} |b|^{2s} \\
&= k'_{s,2s} |\log|b|| \;\; |b|^{2s} + \left(k_{s,2s} + k'_{s,2s}\right) |b|^{2s} \\
&= k'_{s,2s} \left(|b|^{2\varepsilon} |\log|b||\right) |b|^{2(s-\varepsilon)} + \left(k_{s,2s} + k'_{s,2s}\right) |b|^{2(s-\varepsilon)+2\varepsilon} \\
&\leq c_{s,\varepsilon} |b|^{2(s-\varepsilon)},
\end{aligned}$$

where $c_{s,\varepsilon} = k'_{s,2s} \max\limits_{|b| \leq r_G} \left(|b|^{2\varepsilon} |\log|b||\right) + k_{s,2s} + k'_{s,2s}$.

The estimate 4.82 requires that $\eta + \delta_G = s - \varepsilon$ where $\delta_G \geq 0$. But $\eta = s - \frac{1}{2}$ so $\delta_G = \frac{1}{2} - \varepsilon$ where $0 < \varepsilon < \frac{1}{2}$.

Conclusion $r_G = 1$ and regarding the validity of 4.82:

If $0 < s < \theta$, $s \neq 1, 2, 3, \ldots$ then we can choose $\delta_G = s - \frac{1}{2} \lfloor 2s \rfloor$.
If $s = 1, 2, 3, \ldots, \theta - 1$ then 4.82 is valid for all $0 \leq \delta_G < \frac{1}{2}$.

Example 2: The shifted thin-plate splines

We will derive the combinations of η and δ_G for which the shifted thin-plate spline basis functions satisfy the requirements of Theorem 156. Corollary 157 then tells us the orders of convergence of the interpolant are $\eta + \delta_G$, which is an improvement of δ_G. It turns out that we can choose any value of δ_G such that $\delta_G \leq 1/2$.

The shifted thin-plate spline weight and basis functions were studied in Subsection 2.3.2 of Chapter 2. The shifted thin-plate spline weight functions are defined by

$$w(\xi) = \frac{1}{\widetilde{e}(s) \widetilde{K}_{s+d/2}(a|\xi|)} |\xi|^{-2\theta+2s+d}, \quad s > -d/2, \qquad (4.102)$$

where

$$\widetilde{K}_\lambda(t) = t^\lambda K_\lambda(t), \quad t, \lambda \in \mathbb{R}^1,$$

and K_λ is the modified Bessel function of the second kind of order λ. The functions defined by

$$G(x) = \begin{cases} \frac{(-1)^{s+1}}{2}\left(a^2+|x|^2\right)^s \log\left(a^2+|x|^2\right), & s = 1,2,3,\ldots, \\ (-1)^{\lceil s \rceil}\left(a^2+|x|^2\right)^s, & s > -d/2,\ s \neq 1,2,3\ldots, \end{cases} \tag{4.103}$$

where $a > 0$, are shifted thin-plate spline basis functions. In fact:

Theorem 160 *(From Theorem 73)*

1. *w has weight function property W1 w.r.t. the set $\mathcal{A} = \{0\}$.*
2. *w also has weight function properties W2.1 and W3.2 for θ and all $\kappa \geq 0$ iff $-d/2 < s < \theta$.*
3. *If $-d/2 < s < \theta$ then G is a basis function of order θ generated by w.*
4. *If $t > 2s$ there exists a constant c_t such that $|G(x)| \leq c_t (1+|x|)^t$ for all x.*

From Theorem 160 we know that the weight function has property W2 for all $\kappa \geq 0$. Thus $G \in C^\infty$ and

$$D^\beta G \in L^1_{loc} \quad \beta \geq 0.$$

By a cunning trick we can modify the calculations used for the thin-plate spline in the first example. This is done by letting the constant $a \in \mathbb{R}^1_+$ vary and writing $G = G(y)$ where $y = (a, x) \in \mathbb{R}^{d+1}$. Now, like the thin-plate spline basis function, G is a homogeneous function in y of order $2s$ and we can use the nice properties of homogenous functions to bound the derivatives w.r.t. x. Indeed the manipulations are much easier here because there are no singularities.

$\boxed{\text{Case 1 } -d/2 < s < \theta,\ s \neq 1,2,3\ldots}$ Now $\eta = \max\limits_{\kappa \geq 0} \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\} = \theta$ by 4.89. By virtue of homogeneity

$$\left|D^\beta G(x)\right| \leq k_{s,|\beta|}(a+|x|)^{2s-|\beta|}, \quad \beta \geq 0,$$

where $k_{s,n} = \max\limits_{|\alpha|=n} \max\limits_{|x|=1} |D^\alpha G(x)|$ and so

$$\int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(z+tb)\right| dt \leq k_{s,|\beta|}\int_0^1 (a+|z+tb|)^{2s-|\beta|}\, dt.$$

If $|\beta| = 2\eta+1 = 2\theta+1$ then $2s-|\beta| = 2s-2\theta-1 < 0$ and consequently for $\beta \neq 0$

$$\int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(z+tb)\right| dt \leq \left(k_{s,|\beta|}a^{2s-|\beta|}\right),$$

so there is polynomial increase in z as required.
When $|\beta| = 2\eta+1 = 2\theta+1$ and $z = 0$

$$\begin{aligned} |b|^{|\beta|}\int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(tb)\right| dt \leq k_{s,|\beta|}|b|^{|\beta|}\int_0^1 (a+|b|\,t)^{2s-|\beta|}\, dt &= k_{s,|\beta|}|b|^{|\beta|}\int_0^1 \frac{dt}{(a+|b|\,t)^{|\beta|-2s}} \\ &< k_{s,|\beta|}|b|^{|\beta|}\int_0^1 \frac{dt}{a^{|\beta|-2s}} \\ &= \frac{k_{s,|\beta|}}{a^{2(\theta-s)+1}}|b|^{|\beta|} \\ &= \frac{k_{s,2\theta+1}}{a^{2(\theta-s)+1}}|b|^{2\eta+1}, \end{aligned}$$

for all b. But we want a bound with the form $\frac{c_{G,\eta}}{2\sigma}|b|^{2\eta+2\delta_G}$ so $\delta_G = \frac{1}{2}$ when $|b| \leq r_G$ where r_G is arbitrary.

$\boxed{\text{Case 2 } s = 1,2,3,\ldots,\theta-1.}$ By 4.89, $\eta = \max\limits_{\kappa \geq 0} \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\} = \theta$.

Now $G(x) = \frac{(-1)^{s+1}}{2}\left(a^2+|x|^2\right)^s \log\left(a^2+|x|^2\right)$ and since any first order derivative of $\log|\cdot|$ is a homogeneous function of order -1 there exist positive constants $k_{s,|\beta|}$ and $k'_{s,|\beta|}$ such that

$$\left|D^\beta G(x)\right| \leq \left(k_{s,|\beta|}\left|\log|x|\right| + k'_{s,|\beta|}\right)|x|^{2s-|\beta|}, \quad \beta \geq 0.$$

If $|\beta| = 2\eta+1 = 2\theta+1$ then $2s-|\beta| = 2s-2\theta-1 < 0$ and so for $\beta \neq 0$

$$\begin{aligned}
& \int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(z+tb)\right| dt \\
\leq\ & \int_0^1 \left(k_{s,|\beta|}(a+|z+tb|)^{2s-|\beta|}\left|\log(a+|z+tb|)\right| + k'_{s,|\beta|}(a+|z+tb|)^{2s-|\beta|}\right) dt \\
<\ & \int_0^1 \left(k_{s,|\beta|}a^{2s-|\beta|}\left|\log(a+|z+tb|)\right| + k'_{s,|\beta|}a^{2s-|\beta|}\right) dt \\
\leq\ & a^{2s-|\beta|}\max\left\{k_{s,|\beta|}, k'_{s,|\beta|}\right\}\left(1+\int_0^1 \left|\log(a+|z+tb|)\right| dt\right).
\end{aligned} \tag{4.104}$$

But

$$\begin{aligned}
\int_0^1 \left|\log(a+|z+tb|)\right| dt &= \int_0^1 \left|\log a + \log\left(1+\frac{|z+tb|}{a}\right)\right| dt \\
&\leq \int_0^1 |\log a|\, dt + \int_0^1 \log\left(1+\frac{|z+tb|}{a}\right) dt \\
&\leq \int_0^1 |\log a|\, dt + \int_0^1 \log\left(1+\frac{|z|+|b|}{a}\right) dt \\
&= |\log a| + \log\left(1+\frac{|z|+|b|}{a}\right),
\end{aligned} \tag{4.105}$$

which implies the required polynomial increase w.r.t. z. Now substitute 4.105 into 4.104 and set $z=0$ to get

$$|b|^{|\beta|}\int_0^1 (1-t)^{2\eta}\left|\left(D^\beta G\right)(tb)\right| dt \leq a^{2s-|\beta|}\left(1+|\log a|+\log\left(1+\frac{r_G}{a}\right)\right)\max\left\{k_{s,|\beta|}, k'_{s,|\beta|}\right\}|b|^{|\beta|},$$

when $|b| \leq r_G$. We require a bound of the form 4.82 i.e. we need $|\beta| = 2\eta + 2\delta_G$ for some $\delta_G \geq 0$. But $|\beta| = 2\eta+1$ so $\delta_G = \frac{1}{2}$ will suffice.

Conclusion For any $s > 0$ we can choose arbitrary $r_G > 0$ and $\delta_G = \frac{1}{2}$.

5

The Exact smoother and its convergence to the data function

5.1 Introduction

5.1.1 In brief

Section 5.2 Some results from previous chapters.

Section 5.3 Existence and uniqueness of the solution to the Exact smoothing problem. Derivation of several well-known identities. Some properties of the Exact smoother mapping.

Section 5.4 We derive equations which express the reproducing kernel matrix in terms of the basis function matrix.

Section 5.5 Matrices, vectors and bases derived from the semi-Riesz representer.

Section 5.6 Matrix equations are derived for the Exact smoother.

Section 5.7 Estimates the order of pointwise convergence of the Exact smoother to its data function on a bounded region.

5.1.2 In more detail

Given independent data $\left\{x^{(i)}\right\}_{i=1}^{N}$ and dependent data $\{y_i\}_{i=1}^{N}$ the *Exact smoothing problem* involves minimizing the functional $\rho\left|f\right|_{w,\theta}^{2}+\frac{1}{N}\sum_{i=1}^{N}\left|f(x^{(i)})-y_i\right|^{2}$ over the *semi-Hilbert space* X_w^{θ} of continuous functions. Here $\rho>0$ is termed the *smoothing parameter*. The descriptor *Exact* is ours. We first prove the existence and uniqueness of a solution to this problem using the theory of *basis functions* and *semi-Hilbert spaces* generated by *weight functions*. We then derive orders for the pointwise convergence of the smoother to its data function as the density of the data increases. Each function in X_w^{θ} is considered a legitimate data function.

Section 5.2 summarizes some of the theory needed from previous chapters.

The weight function can be considered to have two parameters, the integer *order parameter* $\theta\geq 1$ and the *smoothness parameter* $\kappa\geq 0$. Such a weight function will generate the spaces X_w^{θ} and the basis functions of order θ. We next state the results we require from Chapter 4 where we studied the minimum seminorm interpolation problem. These results were adapted from Light and Wayne. These results start with the fundamental concept of a *unisolvent set* of points. Using a *minimal unisolvent set* we define the Lagrange polynomial interpolation operator $\mathcal{P}$ and $\mathcal{Q}=I-\mathcal{P}$ as well as the unisolvency matrices and the *Light norm* which makes X_w^{θ} a reproducing kernel Hilbert space. Thus the *Riesz representer* R_x of the evaluation functional $f\to f\left(x\right)$ exists and can be expressed in terms of any basis function. In fact, the Riesz representer of the evaluation functional $f\to D^{\gamma}f\left(x\right)$ is $D_x^{\gamma}R_x$ for $\left|\gamma\right|\leq\lfloor\kappa\rfloor$. In this document we follow Light and Wayne and instead of using the reproducing kernel we use the Riesz representer of the evaluation functional $f\to f\left(x\right)$ which can be readily generalized to the evaluation of derivatives.

We next present two types of matrix, the *basis function matrix* and the *reproducing kernel matrix*. The basis function, reproducing kernel and unisolvency matrices will be used to construct the block matrix

equations for the basis function smoother of this document. It will turn out that the Exact smoother will lie in the same *finite dimensional basis function space* $W_{G,X}$ as the minimum seminorm interpolant. Here $X = \{x^{(i)}\}_{i=1}^{N}$ denotes a unisolvent set of independent data points and G is a basis function. A function in $W_{G,X}$ has the form

$$\sum_{i=1}^{N} \phi_i G\left(x - x^{(i)}\right) + p : \phi_i \in \mathbb{C},\ P_X^T \phi = 0,\ p \in P_\theta, \tag{5.1}$$

where P_X is a *unisolvency matrix* and $G\left(x - x^{(i)}\right)$ is termed a *data-translated basis function*. Since the interpolant will be shown to be in $W_{G,X}$ its finite-dimensionality allows the matrix equations to be derived for the ϕ_i and the coefficients of a basis for P_θ. The *vector-valued evaluation operator* $\widetilde{\mathcal{E}}_X f = \left(f\left(x^{(i)}\right)\right)$ is introduced. This operator and its adjoint under the Light norm will be fundamental to solving the Exact smoothing problem.

In Section 5.3 we will use this mathematical machinery to define the Exact smoothing problem and then solve it using the Hilbert space technique of *orthogonal projection* based on the vector-valued evaluation operator $\widetilde{\mathcal{E}}_X$ and its adjoint. The Exact smoothing problem involves minimizing a functional over the space X_w^θ. More specifically, a special Hilbert space is constructed based on the smoothing functional and then the Exact smoothing problem is expressed as one of minimizing the distance between a point generated by the dependent data and a hyperplane generated by the independent data. Standard orthogonal projection results then yield the existence of a unique solution to the Exact smoothing problem which lies in the finite dimensional subspace $W_{G,X}$.

Using the unisolvency matrix and the reproducing kernel matrix a block matrix equation is derived for the values of the Exact smoother on X, assuming the Riesz representer basis $\{R_{x^{(i)}}\}$ for $W_{G,X}$. Then using this equation and relationships between the basis function matrix and the unisolvency matrix a block matrix equation is derived for the Exact smoother, assuming the data-translated basis function basis for $W_{G,X}$. The block matrix is constructed from a basis function matrix and a unisolvency matrix.

The estimates for the order of pointwise convergence will use the Riesz representer and the operators which were used to solve the smoothing problem. We also employ a special lemma derived from the Lagrange interpolation theory from Light and Wayne [10]. The dependent data y is generated using the data functions in the data space X_w^θ so that $y = \widetilde{\mathcal{E}}_X f$ for some $f \in X_w^\theta$. Given a bounded open set Ω, we are interested in the behavior of $\max_{x \in \overline{\Omega}} |u_I(x) - f(x)|$ as some measure h_X of the 'density' of the data points X increases. Following Light and Wayne [10] we have used the measure $h_X = \sup_{\omega \in \Omega} \operatorname{dist}(\omega, X)$. We show in Theorem 191 that if the weight function has order θ and smoothness parameter κ then we have the smoother error estimate

$$|s_e(x) - f_d(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta + \delta_G} + \sqrt{N\rho}\right), \quad x \in \overline{\Omega}, \tag{5.2}$$

where $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$ and some $\delta_G \geq 0$. Here c_G, $c_{\Omega,\theta}$, $K'_{\Omega,\theta}$ are constants independent of X and when $\rho = 0$ we have an estimate for the interpolant derived in Chapter 4. This result is illustrated using the radial thin-plate and surface thin-plate splines.

Unfortunately, for given h_X, $N = |X|$ can be arbitrarily large so to obtain an *order of pointwise convergence* a relationship between N and h_X is obtained by numerically constructing several X in $[-1.5, 1.5] \subset \mathbb{R}^1$ using a uniform statistical distribution and then using a least squares fit to obtain $h_X \simeq 3.09 N^{-0.81}$. Substituting for N in 5.2 and minimizing the right side using ρ we show that given $h_{X_k} \to 0$ there is a sequence $\rho_k \to 0$ such that $\left|s_e^{(k)}(x) - f_d(x)\right| \leq C_1 \left(h_{X_k}\right)^{\eta + \delta_G} \to 0$. We say the order of convergence is at least $\left(h_{X_k}\right)^{\eta + \delta_G}$.

For special data functions that are linear combinations of Riesz representers $R_{x'_i}$ a *doubled order of convergence* of $2\eta + 2\delta_G$ is demonstrated.

I have not included the results of any numerical experiments concerning the smoother error estimates derived in this chapter.

5.2 Preparation

We will require the following results from previous chapters concerning the data spaces X_w^θ, the Riesz representer R_x and the spaces $W_{G,X}$.

5.2.1 The data spaces X_w^θ

Summary 161 *Suppose the weight function w has property W2. If $f \in X_w^\theta$ then $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ and we can define the function $f_F : \mathbb{R}^d \to \mathbb{C}$ a.e. by: $f_F = \widehat{f}$ on $\mathbb{R}^d \backslash 0$. Further:*

1. *The seminorm 1.22 of X_w^θ satisfies (Theorem 25)*
$$\int w\left|\cdot\right|^{2\theta}\left|f_F\right|^2 = \left|f\right|_{w,\theta}^2. \tag{5.3}$$
2. *An alternative definition of X_w^θ is (Theorem 25):*
$$X_w^\theta = \left\{f \in S' : \xi^\alpha \widehat{f} \in L_{loc}^1 \text{ if } \left|\alpha\right| = \theta;\ \int w\left|\cdot\right|^{2\theta}\left|f_F\right|^2 < \infty\right\}. \tag{5.4}$$
Note that the condition $\widehat{f} \in L_{loc}^1\left(\mathbb{R}^d \backslash 0\right)$ is implied by $\xi^\alpha \widehat{f} \in L_{loc}^1$ for all $\left|\alpha\right| = \theta$.
3. *The functional $\left|\cdot\right|_{w,\theta}$ is a seminorm and* null $\left|\cdot\right|_{w,\theta} = P_\theta$. *Also $P \cap X_w^\theta = P_\theta$ (Theorem 25).*
4. *X_w^θ is complete in the seminorm sense for all orders $\theta \geq 1$ (Theorem 37).*
5. *Suppose w also has property W3 for order θ and κ. Then $X_w^\theta \subset C_{BP}^{(\lfloor\kappa\rfloor)}$ (Theorem 40).*

5.2.2 The Riesz representer R_x of $u \to u\left(x\right)$

This summary combines results from Section 4.5.

Summary 162 ***Properties of the Riesz representer** R_x **of the evaluation functional** $u \to u\left(x\right)$ **on** X_w^θ**:***

1. *Given any minimal θ-unisolvent set of points $A = \{a_i\}_{i=1}^M$ we endow the semi-Hilbert space X_w^θ with the Light inner product $\left(\cdot,\cdot\right)_{w,\theta}$ defined using the cardinal basis $\{l_i\}_{i=1}^M$ that corresponds to A. Then the representer is given by 4.35 where G is any basis function of order θ generated by the weight function w with properties W2 and W3 for order θ.*
2. *For all $u \in X_w^\theta$, $u\left(x\right) = \left(u, R_x\right)_{w,\theta}$.*
3. *R_x is unique and is independent of the basis function used to define it.*
4. *$R_x\left(y\right) = \overline{R_y\left(x\right)}$.*
5. *$R_x\left(a_i\right) = l_i\left(x\right)$.*
6. *$\mathcal{Q}u\left(x\right) = \left\langle u, R_x\right\rangle_{w,\theta}$ when $u \in X_w^\theta$.*
7. *$\mathcal{P}R_x = \sum\limits_{i=1}^M l_i\left(x\right) l_i$.*

5.2.3 The basis function spaces $\dot{W}_{G,X}$ and $W_{G,X}$

The importance of the finite dimensional spaces $\dot{W}_{G,X}$ and $W_{G,X}$ is that they contain the solutions to the variational interpolation and smoothing problems studied in this document.

Definition 163 *(Copy of Definition 130)* ***The basis function spaces** $W_{G,X}$ **and** $\dot{W}_{G,X}$*

Suppose the weight function w has properties W2 and W3 for order $\theta \geq 1$ and κ. Then the basis distributions of order θ are continuous functions and we let G be a basis function. Let $X = \left\{x^{(i)}\right\}_{i=1}^N$ be a θ-unisolvent set of distinct points in $\mathbb{R}^d$ and set $M = \dim P_\theta$. Next choose a real-valued basis $\{p_j\}_{j=1}^M$ of P_θ and calculate the unisolvency matrix $P_X = \left(p_j\left(x^{(i)}\right)\right)$. We can now define

$$\dot{W}_{G,X} = \left\{\sum_{i=1}^N v_i G\left(x - x^{(i)}\right) : (v_i) \in \mathbb{C}^N \text{ and } P_X^T v = 0\right\},$$
$$W_{G,X} = \dot{W}_{G,X} + P_\theta.$$

Summary 164 *From Section 4.7 we know that set-wise:*

1. $W_{G,X}$ *is independent of the basis function of order* θ *used to define it (Theorem 132).*

2. $\dot{W}_{G,X}$ *and* $W_{G,X}$ *are independent of the basis used to define* P_θ *(Theorem 131).*

3. $W_{G,X}$ *is independent of the ordering of the points in* X *(Theorem 132).*

The next theorem restates more results from Section 4.7.

Remark 165 ***Summary 166*** *The spaces* $\dot{W}_{G,X}$ *and* $W_{G,X}$ *have the following properties: from Theorem 134:*

1. *If* $f_v(x) = \sum\limits_{k=1}^{N} v_k G\left(x - x^{(k)}\right)$ *then* $|f_v|^2_{w,\theta} = (2\pi)^{\frac{d}{2}} v^T G_{X,X} \overline{v}$.

2. $G_{X,X}$ *is conditionally positive definite on* $\operatorname{null} P_X^T$ *i.e. when* $P_X^T v = 0$ *and* $v \neq 0$ *we have* $v^T G_{X,X} \overline{v} > 0$.

3. $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$, $\dim \dot{W}_{G,X} = N - M$ *and* $\dim W_{G,X} = N$.

4. $X_w^\theta = W_{G,X} \oplus W_{G,X}^\perp$ *where*

$$W_{G,X}^\perp = \left\{ u \in X_w^\theta : u\left(x^{(k)}\right) = 0 \text{ for all } x^{(k)} \in X \right\}.$$

   From Corollary 135:

5. *If* $\{p_j\}_{j=1}^M$ *is basis for* P_θ *then the representation*

$$W_{G,X} = \left\{ \sum_{i=1}^{N} \alpha_i G\left(\cdot - x^{(i)}\right) + \sum_{j=1}^{M} \beta_j p_j : P_X^T \alpha = 0,\ \alpha = (\alpha_i),\ \alpha_i, \beta_j \in \mathbb{C} \right\}, \tag{5.5}$$

   is unique in terms of α_i *and* β_j.

   From Corollary 137:

6. *Suppose* $A \subset X$ *is a minimal unisolvent set and suppose* R_x *is the Riesz representer of the functional* $f \to f(x)$ *w.r.t. the Light norm, both being defined using* A. *Then* $\dot{W}_{G,X}$ *has basis* $\left\{R_{x^{(i)}} : x^{(i)} \notin A\right\}$ *and* $W_{G,X}$ *has basis* $\{R_{x^{(i)}}\}_{i=1}^N$.

5.3 Solution of the Exact smoothing problem and its properties

The Exact smoothing problem involves minimizing a functional over the space of continuous functions X_w^θ. The functional is constructed using a smoothing parameter $\rho > 0$, the seminorm $|\cdot|_{w,\theta}$ of X_w^θ and a set of distinct, scattered data points $\left\{\left(x^{(i)}, y_i\right)\right\}_{i=1}^N$, $x^{(i)} \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$. As with the interpolation problem we require the $x^{(i)}$ to be distinct. The data will sometimes be specified using the notation $[X, y]$ where $X = \left\{x^{(i)}\right\}_{i=1}^N$ is termed the independent data and $y = \{y_i\}_{i=1}^N$ is termed the dependent data. The functional to be minimized is

$$J_e[f] = \rho |f|^2_{w,\theta} + \frac{1}{N} \sum_{i=1}^{N} \left| f\left(x^{(i)}\right) - y_i \right|^2, \quad f \in X_w^\theta. \tag{5.6}$$

The Exact smoothing problem is now stated as:

Definition 167 ***The Exact smoothing problem and the Exact smoother***

$\boxed{\textit{Minimize the Exact smoothing functional } J_e \textit{ on the space } X_w^\theta.}$

A solution to this problem will be called an ***Exact smoother*** *of the data* $[X, y]$.

The first component $\rho\,|f|_{w,\theta}^2$ of the functional can be regarded as a global smoothing component and the second component as the localizing least squares component.

It is clear that when $\rho = 0$ any interpolant of the data minimizes the Exact smoother functional. Thus, from part 4 Theorem 166 the set of solutions is $u_I + W_{G,X}^{\perp}$ where u_I is the minimal seminorm interpolant studied in Chapter 4. In Remark 184 we will note that as $\rho \to 0$ the matrix equation representing the solution to this problem becomes that of the minimal norm interpolation problem so the limit "seeks out" one particular interpolant.

5.3.1 Existence, uniqueness and formulas for the smoother

Using the technique of orthogonal projection it will be shown below that a solution to the smoothing problem exists and is unique. The proof will be carried out within a Hilbert space framework by formulating the smoothing functional in terms of a special inner product on the Hilbert product space $V = X_w^\theta \otimes \mathbb{C}^N$. To this end I will introduce the following definitions:

Definition 168 *For order* $\theta \geq 1$*:*

1. *Suppose* $\rho > 0$. *Let* $V = X_w^\theta \otimes \mathbb{C}^N$ *be the Hilbert product space with inner product*

$$((u_1, \widetilde{u}_2), (v_1, \widetilde{v}_2))_V = \rho \langle u_1, v_1 \rangle_{w,\theta} + \frac{1}{N} (\widetilde{u}_2, \widetilde{v}_2)_{\mathbb{C}^N},$$

and $\|f\|_V^2 = (f, f)_V$.

2. *Let* $\mathcal{L}_X : X_w^\theta \to V$ *be defined by* $\mathcal{L}_X f = \left(f, \widetilde{\mathcal{E}}_X f\right)$ *where* $\widetilde{\mathcal{E}}_X$ *is the evaluation operator of Definition 138.*

Remark 169 *The smoothing functional can be expressed in terms of* $\mathcal{L}_X$ *and the data as follows: set* $\varsigma = (0, y) \in V$ *where* $y = (y_i)$ *is the dependent data given in the Exact smoothing problem. Then for* $f \in X_w^\theta$

$$\begin{aligned}\|\mathcal{L}_X f - \varsigma\|_V^2 = \left\|\left(f, \widetilde{\mathcal{E}}_X f\right) - (0, y)\right\|_V^2 = \left\|\left(f, \widetilde{\mathcal{E}}_X f - y\right)\right\|_V^2 &= \rho\,|f|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X f - y\right|^2 \\ &= J_e[f].\end{aligned}$$

The operator $\mathcal{L}_X : X_w^\theta \to V$ and its adjoint $\mathcal{L}_X^* : V \to X_w^\theta$ have the following properties:

Theorem 170 *Suppose* X *is a* θ*-unisolvent set and* $A \subset X$ *is a minimal unisolvent set. Use* A *to define the Light norm* $\|\cdot\|_{w,\theta}$ *and the Lagrange interpolation operator* $\mathcal{P}$. *Then when* $\rho > 0$*:*

1. $\|\mathcal{L}_X f\|_V$ *and* $\|f\|_{w,\theta}$ *are equivalent norms on* X_w^θ.

2. $\mathcal{L}_X : X_w^\theta \to V$ *is continuous and 1-1.*

3. $\mathcal{L}_X^* : V \to X_w^\theta$ *is continuous w.r.t. the Light norm and if* $u = (u_1, \widetilde{u}_2) \in V$ *then*

$$\mathcal{L}_X^* u = \rho \mathcal{Q} u_1 + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{u}_2,$$

with $\operatorname{range} \mathcal{L}_X^* = X_w^\theta$.

4. $\mathcal{L}_X^* \mathcal{L}_X : X_w^\theta \to X_w^\theta$ *and*

$$\mathcal{L}_X^* \mathcal{L}_X = \rho \mathcal{Q} + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X.$$

5. $\mathcal{L}_X^* \mathcal{L}_X$ *is 1-1 on* X_w^θ. *Also,* $(\mathcal{L}_X^* \mathcal{L}_X)^{-1}$ *is continuous on* $\operatorname{range} \mathcal{L}_X^* \mathcal{L}_X$.

Proof. Part 1 From the definition of $\mathcal{L}_X$, if $f \in X_w^\theta$

$$\|\mathcal{L}_X f\|_V^2 = \left\|\left(f, \widetilde{\mathcal{E}}_X f\right)\right\|_V^2 = \rho\,|f|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X f\right|_{\mathbb{C}^N}^2 \leq const\,\|f\|_{w,\theta}^2,$$

since $\widetilde{\mathcal{E}}_X : X_w^\theta \to \mathbb{R}^N$ is continuous by part 1 of Theorem 139. Also

$$\begin{aligned}\|f\|_{w,\theta}^2 = |f|_{w,\theta}^2 + \left|\widetilde{\mathcal{E}}_A f\right|_{\mathbb{C}^M}^2 &\leq \left(\min\left\{\rho, \frac{1}{N}\right\}\right)^{-1}\left(\rho\,|f|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X f\right|_{\mathbb{C}^N}^2\right)\\ &= \left(\min\left\{\rho, \frac{1}{N}\right\}\right)^{-1}\|\mathcal{L}_X f\|_V^2,\end{aligned}$$

since $\rho > 0$.

Part 2 Part 1 implies $\mathcal{L}_X$ is continuous and that $\mathcal{L}_X$ is 1-1 is clear from its definition.

Part 3 Since $\mathcal{L}_X$ is a continuous operator the adjoint $\mathcal{L}_X^* : V \to X_w^\theta$ exists and is continuous. Further, if $f \in X_w^\theta$ and $u = (u_1, \widetilde{u}_2) \in V$ then by the properties of the Light norm given in Theorem 116

$$\begin{aligned}(\mathcal{L}_X f, u)_V = \left(\left(f, \widetilde{\mathcal{E}}_X f\right), (u_1, \widetilde{u}_2)\right)_V &= \rho\,\langle f, u_1\rangle_{w,\theta} + \frac{1}{N}\left(\widetilde{\mathcal{E}}_X f, \widetilde{u}_2\right)_{\mathbb{R}^N}\\ &= \rho\,(f, \mathcal{Q}u_1)_{w,\theta} + \frac{1}{N}\left(f, \widetilde{\mathcal{E}}_X^* \widetilde{u}_2\right)_{w,\theta}\\ &= \left(f, \rho\mathcal{Q}u_1 + \frac{1}{N}\widetilde{\mathcal{E}}_X^* \widetilde{u}_2\right)_{w,\theta}.\end{aligned}$$

Hence $\mathcal{L}_X^* u = \rho\mathcal{Q}u_1 + \frac{1}{N}\widetilde{\mathcal{E}}_X^* \widetilde{u}_2$.

Part 4

$$\mathcal{L}_X^*\mathcal{L}_X f = \rho\mathcal{Q}\left((\mathcal{L}_X f)_1\right) + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\left((\mathcal{L}_X f)_2\right) = \rho\mathcal{Q}f + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f.$$

Part 5 Suppose $\mathcal{L}_X^*\mathcal{L}_X f = 0$. Then $0 = (\mathcal{L}_X^*\mathcal{L}_X f, f)_{w,\theta} = (\mathcal{L}_X f, \mathcal{L}_X f)_{w,\theta} = \|\mathcal{L}_X f\|_{w,\theta}^2$ so that $\mathcal{L}_X f = 0$ and $f = 0$ since $\mathcal{L}_X$ is 1-1. Hence $\mathcal{L}_X^*\mathcal{L}_X$ is 1-1.

Since $\mathcal{L}_X^*\mathcal{L}_X$ is continuous it has closed range. By the bounded inverse theorem $(\mathcal{L}_X^*\mathcal{L}_X)^{-1}$ is also continuous on range $\mathcal{L}_X^*\mathcal{L}_X$. ∎

Using the Hilbert space technique of orthogonal projection the next theorem shows that when $\rho > 0$ the Exact smoothing problem of Definition 167 has a unique solution in X_w^θ.

Theorem 171 *Fix $y \in \mathbb{R}^N$ and let $\varsigma = (0, y) \in V$. Then for $\rho > 0$ there exists a unique function $s_e \in X_w^\theta$ which solves the Exact smoothing problem with data $[X, y]$. This solution has the following properties:*

1. $\|\mathcal{L}_X s_e - \varsigma\|_V < \|\mathcal{L}_X f - \varsigma\|_V$ *for all* $f \in X_w^\theta - \{s_e\}$.
2. $(\mathcal{L}_X s_e - \varsigma, \mathcal{L}_X s_e - \mathcal{L}_X f)_V = 0$ *for all* $f \in X_w^\theta$.
3. $\|\mathcal{L}_X s_e - \varsigma\|_V^2 + \|\mathcal{L}_X s_e - \mathcal{L}_X f\|_V^2 = \|\mathcal{L}_X f - \varsigma\|_V^2$ *for all* $f \in X_w^\theta$.
   This equality is equivalent to that of part 2.
4. $s_e = \frac{1}{N}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\widetilde{\mathcal{E}}_X^* y$.
5. *Since $\theta \geq 1$, $p \in P_\theta$ implies $p = \frac{1}{N}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X p$. We say the Exact smoother preserves polynomials up to order θ.*

Proof. Parts 1,2,3 Since $\mathcal{L}_X$ is continuous, we know that the hyperspace $\mathcal{L}_X\left(X_w^\theta\right)$ is closed and hence that the translated hyperspace $\mathcal{L}_X\left(X_w^\theta\right) - \varsigma$ is also closed. But from the remark following Definition 168 $J_e[f] = \|\mathcal{L}_X f - \varsigma\|_V^2$ and the Exact smoothing problem becomes minimize $\|\mathcal{L}_X f - \varsigma\|_V$ over X_w^θ. Thus by a well-known orthogonal projection result concerning the distance between a point and a closed subspace, there exists a unique element of $\mathcal{L}_X(X_w^\theta) - \varsigma$, call it $\mathcal{L}_X s_e - \varsigma$ such that s_e satisfies parts 1, 2 and 3 of this theorem.

Part 4 Using the equation proved in part 2

$$0 = (\mathcal{L}_X s_e - \varsigma, \mathcal{L}_X s_e - \mathcal{L}_X f)_V = (\mathcal{L}_X s_e - \varsigma, \mathcal{L}_X (s_e - f))_V = (\mathcal{L}_X^* (\mathcal{L}_X s_e - \varsigma), s_e - f)_{w,\theta},$$

for all $f \in X_w^\theta$. Thus

$$\mathcal{L}_X^* (\mathcal{L}_X s_e - \varsigma) = 0, \tag{5.7}$$

and therefore

$$\mathcal{L}_X^* \mathcal{L}_X s_e = \mathcal{L}_X^* \varsigma = \mathcal{L}_X^* (0, y) = \frac{1}{N} \widetilde{\mathcal{E}}_X^* y.$$

But by part 5 of Theorem 170, $\mathcal{L}_X^* \mathcal{L}_X$ is one-to-one and so $s_e = \frac{1}{N} (\mathcal{L}_X^* \mathcal{L}_X)^{-1} \widetilde{\mathcal{E}}_X^* y$.

Part 5 Substituting $y = \widetilde{\mathcal{E}}_X p$ into the result proved in the previous part gives $s_e = \frac{1}{N} (\mathcal{L}_X^* \mathcal{L}_X)^{-1} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X p$, or equivalently $N \mathcal{L}_X^* \mathcal{L}_X s_e = \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X p$. But from part 4 of Theorem 170, $\mathcal{L}_X^* \mathcal{L}_X - \rho \mathcal{Q} = \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X$.

Substituting this as an expression for $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X$ into the last equation we get

$$N \mathcal{L}_X^* \mathcal{L}_X s_e = N (\mathcal{L}_X^* \mathcal{L}_X p - \rho \mathcal{Q} p) = N \mathcal{L}_X^* \mathcal{L}_X p,$$

since $\mathcal{Q} p = (I - \mathcal{P}) p = 0$. Finally, by part 5 of Theorem 170, $\mathcal{L}_X^* \mathcal{L}_X$ is one-to-one and hence $s_e = p$.
■

5.3.2 *Various identities*

The last theorem showed that the Exact smoothing problem has a unique minimizer in X_w^θ. In the next corollary we prove that the smoother lies in the space $W_{G,X}$ introduced in Subsection 5.2.3. In the next two corollaries we will also prove several well-known identities which involve all the data points. These identities relate the Hilbert space properties and the pointwise properties of the data and the Exact smoother.

Corollary 172 *Suppose $X = \{x^{(k)}\}_{k=1}^N$ is θ-unisolvent so that X contains a minimal unisolvent subset A. Use this subset to define the operators $\mathcal{P}$, $\mathcal{Q}$ and the Light norm. Then the unique solution $s_e \in X_w^\theta$ of the Exact smoothing problem with data $[X, y]$ has the following properties:*

1. *$s_e \in W_{G,X}$ and $s_e = \mathcal{P} s_e - \frac{1}{N\rho} \widetilde{\mathcal{E}}_X^* \left(\widetilde{\mathcal{E}}_X s_e - y \right)$.*
2. *For all $f \in X_w^\theta$ we have*

$$\rho |s_e|_{w,\theta}^2 + \frac{1}{N} \sum_{k=1}^N \left| s_e\left(x^{(k)}\right) - y_k \right|^2 + \rho |s_e - f|_{w,\theta}^2 + \frac{1}{N} \sum_{k=1}^N \left| s_e\left(x^{(k)}\right) - f\left(x^{(k)}\right) \right|^2 \tag{5.8}$$

$$= \rho |f|_{w,\theta}^2 + \frac{1}{N} \sum_{k=1}^N \left| f\left(x^{(k)}\right) - y_k \right|^2,$$

or

$$J_e[s_e] + \rho |s_e - f|_{w,\theta}^2 + \frac{1}{N} \sum_{k=1}^N \left| s_e\left(x^{(k)}\right) - f\left(x^{(k)}\right) \right|^2 = J_e[f]. \tag{5.9}$$

3. *$P_X^T ((s_e)_X - y) = 0$.*

Proof. Part 1

$$0 = \mathcal{L}_X^* (\mathcal{L}_X s_e - \varsigma) = \mathcal{L}_X^* \mathcal{L}_X s_e - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y = \rho \mathcal{Q} s_e + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X s_e - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y$$

$$= \rho s_e - \rho \mathcal{P} s_e + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \left(\widetilde{\mathcal{E}}_X s_e - y \right),$$

so that

$$s_e = \mathcal{P} s_e - \frac{1}{N\rho} \widetilde{\mathcal{E}}_X^* \left(\widetilde{\mathcal{E}}_X s_e - y \right),$$

and $s_e \in W_{G,X}$ by part 2 of Theorem 139.

Part 2 By part 3 of Theorem 171

$$\|\mathcal{L}_X s_e - \varsigma\|_V^2 + \|\mathcal{L}_X s_e - \mathcal{L}_X f\|_V^2 = \|\mathcal{L}_X f - \varsigma\|_V^2, \quad f \in X_w^\theta.$$

Using Remark 169 and the definition of $\mathcal{L}_X$ this equation becomes

$$\rho\left|s_e\right|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X s_e - y\right|^2 + \rho\left|s_e - f\right|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X\left(s_e - f\right)\right|^2 = \rho\left|f\right|_{w,\theta}^2 + \frac{1}{N}\left|\widetilde{\mathcal{E}}_X f - y\right|^2,$$

for all $f \in X_w^\theta$ and $y \in \mathbb{R}^N$, which proves this part.

Part 3 If $p \in P_\theta$, then from part 1

$$\begin{aligned}\frac{1}{N\rho}\left(\widetilde{\mathcal{E}}_X^*\left(\left(s_e\right)_X - y\right), p\right)_{w,\theta} = \frac{1}{N\rho}\left(\widetilde{\mathcal{E}}_X^*\left(\widetilde{\mathcal{E}}_X s_e - y\right), p\right)_{w,\theta} &= \left(\mathcal{P}s_e - s_e, p\right)_{w,\theta} \\ &= -\left(\mathcal{Q}s_e, p\right)_{w,\theta} \\ &= -\left\langle s_e, p\right\rangle_{w,\theta} \\ &= 0.\end{aligned}$$

But

$$0 = \left(\widetilde{\mathcal{E}}_X^*\left(\left(s_e\right)_X - y\right), p\right)_{w,\theta} = \left(\left(s_e\right)_X - y, \widetilde{\mathcal{E}}_X p\right) = \sum_{i=1}^N \left(\left(s_e\right)_X - y\right)_i \overline{p\left(x^{(i)}\right)},$$

so by part 2 of Theorem 104, $P_X^T\left(\left(s_e\right)_X - y\right) = 0$. ■

We next prove some standard results which express the smoothing functional and the seminorm of the smoother in terms of the dependent data and the values taken by the smoother on the independent data.

Corollary 173 *Suppose $\left\{\left(x^{(k)}, y_k\right)\right\}_{k=1}^N$ is the data to be smoothed. Then the Exact smoothing function s_e has the following properties:*

1. *For all $p \in P_\theta$*

$$2\rho\left|s_e\right|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^N\left|s_e\left(x^{(k)}\right) - y_k\right|^2 + \frac{1}{N}\sum_{k=1}^N\left|s_e\left(x^{(k)}\right) - p\left(x^{(k)}\right)\right|^2 = \frac{1}{N}\sum_{k=1}^N\left|p\left(x^{(k)}\right) - y_k\right|^2.$$

2. $2\rho\left|s_e\right|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^N\left|s_e\left(x^{(k)}\right) - y_k\right|^2 + \frac{1}{N}\sum_{k=1}^N\left|s_e\left(x^{(k)}\right)\right|^2 = \frac{1}{N}\sum_{k=1}^N\left|y_k\right|^2.$

3. $\left|s_e\right|_{w,\theta}^2 = \frac{1}{N\rho}\operatorname{Re}\sum_{k=1}^N \overline{s_e\left(x^{(k)}\right)}\left(y_k - s_e\left(x^{(k)}\right)\right).$

4. $J_e\left[s_e\right] = \frac{1}{N}\operatorname{Re}\sum_{k=1}^N\left(y_k - s_e\left(x^{(k)}\right)\right)\overline{y_k}.$

Proof. Part 1 Substitute $f = p \in P_\theta$ in the equation of part 2 of Corollary 172 and use the fact that $\left|g + p\right|_{w,\theta} = \left|g\right|_{w,\theta}$ when $g \in X_w^\theta$.

Part 2 Substitute $f = 0$ in the equation of part 2 of Corollary 172.

Part 3 Substitute the expansion

$$\left|s_e\left(x^{(k)}\right) - y_k\right|^2 = \left|s_e\left(x^{(k)}\right)\right|^2 - 2\operatorname{Re} s_e\left(x^{(k)}\right)\overline{y_k} + \left|y_k\right|^2,$$

in part 2 of Corollary 172 so that

$$\begin{aligned}
2\rho\left|s_e\right|_{w,\theta}^2 &= \frac{1}{N}\sum_{k=1}^{N}\left|y_k\right|^2 - \frac{1}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right)-y_k\right|^2 - \frac{1}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right)\right|^2 \\
&= \frac{1}{N}\sum_{k=1}^{N}\left|y_k\right|^2 - \frac{1}{N}\sum_{k=1}^{N}\left(\left|s_e\left(x^{(k)}\right)\right|^2 - 2\operatorname{Re} s_e\left(x^{(k)}\right)\overline{y_k} + \left|y_k\right|^2\right) \\
&\qquad -\frac{1}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right)\right|^2 \\
&= \frac{2}{N}\sum_{k=1}^{N}\operatorname{Re} s_e\left(x^{(k)}\right)\overline{y_k} - \frac{2}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right)\right|^2 \\
&= \frac{2}{N}\operatorname{Re}\sum_{k=1}^{N}\left(s_e\left(x^{(k)}\right)\overline{y_k} - s_e\left(x^{(k)}\right)\overline{s_e\left(x^{(k)}\right)}\right) \\
&= \frac{2}{N}\operatorname{Re}\sum_{k=1}^{N} s_e\left(x^{(k)}\right)\left(\overline{y_k} - \overline{s_e\left(x^{(k)}\right)}\right) \\
&= \frac{2}{N}\operatorname{Re}\sum_{k=1}^{N}\overline{s_e\left(x^{(k)}\right)}\left(y_k - s_e\left(x^{(k)}\right)\right).
\end{aligned}$$

Part 4 Substitute the formula for $\left|s_e\right|_{w,\theta}^2$ of part 3 into the definition of $J_e\left[s_e\right]$ to get

$$\begin{aligned}
J_e\left[s_e\right] &= \rho\left|s_e\right|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right)-y_k\right|^2 \\
&= \frac{1}{N}\operatorname{Re}\sum_{k=1}^{N} s_e\left(x^{(k)}\right)\left(\overline{y_k} - \overline{s_e\left(x^{(k)}\right)}\right) + \\
&\quad + \frac{1}{N}\sum_{k=1}^{N}\left(\left|s_e\left(x^{(k)}\right)\right|^2 - 2\operatorname{Re} s_e\left(x^{(k)}\right)\overline{y_k} + \left|y_k\right|^2\right) \\
&= \frac{1}{N}\sum_{k=1}^{N}\left|y_k\right|^2 - \frac{1}{N}\operatorname{Re}\sum_{k=1}^{N} s_e\left(x^{(k)}\right)\overline{y_k} \\
&= \frac{1}{N}\operatorname{Re}\sum_{k=1}^{N}\left(y_k - s_e\left(x^{(k)}\right)\right)\overline{y_k}.
\end{aligned}$$

■

5.3.3 Data functions and the Exact smoother mapping

Part 4 of Theorem 171 allows us to define the mapping between a data function and its corresponding Exact smoother. We call this the Exact smoother mapping:

Definition 174 ***Data functions and the Exact smoother mapping***
Given an independent data set X, we shall assume that each member of X_w^θ can act as a legitimate data function f and generate the data vector $\widetilde{\mathcal{E}}_X f$. By part 5 of Theorem 170 the linear operator $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}$ exists and is continuous. Thus part 4 of Theorem 171 enables us to define the continuous linear mapping $\mathcal{S}_X^e : X_w^\theta \to W_{G,X}$ from the data functions to the corresponding unique Exact smoother by

$$\mathcal{S}_X^e f = \frac{1}{N}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f, \quad f \in X_w^\theta. \tag{5.10}$$

We will now prove some properties of the operator $\mathcal{L}_X^*\mathcal{L}_X$ which in turn will be used to prove some basic properties of the Exact smoother mapping.

Theorem 175 ***Properties of $\mathcal{L}_X^*\mathcal{L}_X$ and $\mathcal{S}_X^e$***

When $\rho > 0$ the composition $\mathcal{L}_X^*\mathcal{L}_X : X_w^\theta \to X_w^\theta$ is a homeomorphism and $\mathcal{L}_X^*\mathcal{L}_X : W_{G,X} \to W_{G,X}$ is also a homeomorphism. Regarding the properties of $\mathcal{S}_X^e$:

$$\mathcal{S}_X^e f = f - \rho \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} \mathcal{Q}f, \qquad f \in X_w^\theta, \tag{5.11}$$

and the operator $\mathcal{S}_X^e$ is continuous. Further:

1. *$|\mathcal{S}_X^e f|_{w,\theta} \leq |f|_{w,\theta}$ and $|(I - \mathcal{S}_X^e) f|_{w,\theta} \leq |f|_{w,\theta}$ when $f \in X_w^\theta$ i.e. $\mathcal{S}_X^e$ and $I - \mathcal{S}_X^e$ are contractions in the seminorm sense.*
2. *$\mathcal{S}_X^e f = f$ iff $f \in P_\theta$.*
3. *$\mathcal{S}_X^e : X_w^\theta \to W_{G,X}$ is onto and* $\operatorname{null} \mathcal{S}_X^e = W_{G,X}^\perp$.
4. *The adjoint of the Exact smoother w.r.t. the Light norm is given by:*

$$\left(\mathcal{S}_X^e\right)^* g = \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} g = g - \rho\mathcal{Q}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} g, \quad g \in X_w^\theta.$$

5. *$\mathcal{S}_X^e$ is self-adjoint iff X is a minimal unisolvent set iff $\mathcal{S}_X^e = \mathcal{P}$ i.e. the Lagrange polynomial interpolation function 4.1.*

Proof. Suppose $f \in X_w^\theta$ and let $s_e = \mathcal{S}_X^e f$. From part 5 of Theorem 170 is is known that $\mathcal{L}_X^*\mathcal{L}_X$ is 1-1. From part 4 of Theorem 171, $\mathcal{L}_X^*\mathcal{L}_X s_e = \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f$ and from part 4 of Theorem 170, $\mathcal{L}_X^*\mathcal{L}_X f = \rho\mathcal{Q}f + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f$ so that $\mathcal{L}_X^*\mathcal{L}_X f = \rho\mathcal{Q}f + \mathcal{L}_X^*\mathcal{L}_X s_e$. Clearly this equation implies equation 5.11 and also

$$\mathcal{Q}f = \mathcal{L}_X^*\mathcal{L}_X \left(\frac{f - s_e}{\rho}\right), \qquad f \in X_w^\theta. \tag{5.12}$$

This last equation proves $\mathcal{Q}\left(X_w^\theta\right) \subset \mathcal{L}_X^*\mathcal{L}_X\left(X_w^\theta\right)$ and if we can show that $P_\theta \subset \mathcal{L}_X^*\mathcal{L}_X\left(X_w^\theta\right)$ then it follows that $\mathcal{L}_X^*\mathcal{L}_X\left(X_w^\theta\right) = X_w^\theta$ and so by the *open mapping* theorem $\mathcal{L}_X^*\mathcal{L}_X$ is a homeomorphism. Suppose $p \in P_\theta$ and for some $f \in X_w^\theta$

$$p = \mathcal{L}_X^*\mathcal{L}_X f = \rho\mathcal{Q}f + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f. \tag{5.13}$$

By Theorem 139, $\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X : X_w^\theta \to W_{G,X}$ is onto and so it follows that a solution g to $\frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X g = p$ exists. Now try a solution to 5.13 of the form $f = g + u$ so that $p = \mathcal{L}_X^*\mathcal{L}_X(g + u)$ and u must satisfy $\mathcal{L}_X^*\mathcal{L}_X u = p - \mathcal{L}_X^*\mathcal{L}_X g = p - \left(\rho\mathcal{Q}g + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X g\right) = -\rho\mathcal{Q}g \in \mathcal{Q}\left(X_w^\theta\right)$. But we already know that $\mathcal{Q}\left(X_w^\theta\right) \subset \mathcal{L}_X^*\mathcal{L}_X\left(X_w^\theta\right)$ so a there exists u such that $\mathcal{L}_X^*\mathcal{L}_X u = -\rho\mathcal{Q}g$ and hence a solution $f = g + u$ to 5.13. Consequently $P_\theta \subset \mathcal{L}_X^*\mathcal{L}_X\left(X_w^\theta\right)$ and $\mathcal{L}_X^*\mathcal{L}_X : X_w^\theta \to X_w^\theta$ is a homeomorphism.

By part 3 of Theorem 166, $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$ and hence $\mathcal{P}\left(W_{G,X}\right) = P_\theta$, $W_{G,X} = \mathcal{Q}\left(W_{G,X}\right) \oplus P_\theta$ and $\mathcal{Q}\left(W_{G,X}\right) = \dot{W}_{G,X}$. Thus 1.60 implies $\dot{W}_{G,X} = \mathcal{Q}\left(W_{G,X}\right) \subset \mathcal{L}_X^*\mathcal{L}_X\left(W_{G,X}\right)$ and so $\mathcal{L}_X^*\mathcal{L}_X\left(W_{G,X}\right) = W_{G,X}$. The *open mapping* theorem then implies that $\mathcal{L}_X^*\mathcal{L}_X : W_{G,X} \to W_{G,X}$ is a homeomorphism.

Part 1 In equation 5.8 let f be the data function and note that the data function interpolates the data.

Part 2 Since $\mathcal{L}_X^*\mathcal{L}_X$ is a homeomorphism, from 5.11 $\mathcal{S}_X^e f = f$ iff $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{Q}f = 0$ iff $\mathcal{Q}f = 0$ iff $f \in P_\theta$, where the last implication follows from part 4 of Theorem 102.

Part 3 By part 6 of Theorem 139, $\mathcal{S}_X^e f = 0$ iff $\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f = 0$ iff $f \in W_{G,X}^\perp$.

From 5.10, $\mathcal{S}_X^e f = \frac{1}{N}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X f$ and since we know from earlier in the proof that $\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X : X_w^\theta \to W_{G,X}$ and $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} : W_{G,X} \to W_{G,X}$ are both onto we have our result.

Part 4 A standard Hilbert space result is that if $\mathcal{K}$ is a homeomorphism then $\left(\mathcal{K}^{-1}\right)^* = \left(\mathcal{K}^*\right)^{-1}$. Thus, since $\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X$ and $\mathcal{L}_X^*\mathcal{L}_X$ are self-adjoint and by Theorem 116 $\mathcal{Q}$ is self-adjoint, taking the adjoint of equations 5.11 and 5.12 easily yields the claimed formulas for $\left(\mathcal{S}_X^e\right)^*$.

Part 5 Starting with the formulas of 5.11 and part 4 we have the equivalences:

$$\begin{aligned} \left(\mathcal{S}_X^e\right)^* = \mathcal{S}_X^e \quad &\text{iff} \quad \mathcal{Q}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} f = \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{Q}f, \quad f \in X_w^\theta, \\ &\text{iff} \quad \mathcal{P}\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} f = \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{P}f, \quad f \in X_w^\theta. \end{aligned}$$

Since $\mathcal{L}_X^* \mathcal{L}_X : X_w^\theta \to X_w^\theta$ is a homeomorphism, if we set $g = (\mathcal{L}_X^* \mathcal{L}_X)^{-1} f$ then

$$(\mathcal{S}_X^e)^* = \mathcal{S}_X^e \text{ iff } \mathcal{L}_X^* \mathcal{L}_X \mathcal{P} g = \mathcal{P} \mathcal{L}_X^* \mathcal{L}_X g, \quad g \in X_w^\theta.$$

Further, by Theorem 170, $\mathcal{L}_X^* \mathcal{L}_X = \rho \mathcal{Q} + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X$ and since $\mathcal{P}\mathcal{Q} = 0$

$$(\mathcal{S}_X^e)^* = \mathcal{S}_X^e \text{ iff } \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X \mathcal{P} g = \mathcal{P} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X g, \quad g \in X_w^\theta.$$

Recall the properties of the Riesz representer R_x given in Subsection 4.5.3. Next suppose $X = \{x^{(k)}\}_{k=1}^N$ is not minimally unisolvent and that the minimal unisolvent set $A \subset X$ was used to construct the smoother, $\mathcal{P}$, $\mathcal{Q}$ etc. Then $\mathcal{P} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X g \in P_\theta$ and so $\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X \mathcal{P} g = \mathcal{P} g + \sum\limits_{x^{(i)} \in X \backslash A} (\mathcal{P} g)\left(x^{(i)}\right) R_{x^{(i)}} \in P_\theta$ with the independence of the functions $\{R_{x^{(i)}}\}_{i=1}^N$ implying that $(\mathcal{P} g)\left(x^{(i)}\right) = 0$ when $x^{(i)} \in X \backslash A$ and $g \in X_w^\theta$. Therefore, given $x^{(i)} \in X \backslash A$ we have $p\left(x^{(i)}\right) = 0$ when $p \in P_\theta$ which is a contradiction. We conclude that $(\mathcal{S}_X^e)^* \neq \mathcal{S}_X^e$ or X is minimally unisolvent.

If X is a minimal unisolvent set and f is a data function then $\mathcal{P} f\left(x^{(k)}\right) = f\left(x^{(k)}\right)$ and $\mathcal{P} f \in P_\theta$ so that the Exact smoother functional satisfies $J_e[\mathcal{P} f] = 0$ and thus $\mathcal{S}_X^e f = \mathcal{P} f$ i.e. $\mathcal{S}_X^e = \mathcal{P}$. Finally, if $\mathcal{S}_X^e = \mathcal{P}$ then $\mathcal{S}_X^e$ is self-adjoint w.r.t. the Light norm by Theorem 116. ■

5.4 Expressing $R_{X,X}$ in terms of $G_{X,X}$

In this section we will derive equations expressing the reproducing kernel matrix in terms of the basis function matrix. I start with some notation:

Definition 176 ***Matrices and vectors derived from the Riesz representer*** R_x ***and the basis function*** G.

If $X = \{x^{(k)}\} \subset \mathbb{R}^d$ *and* $y \in \mathbb{R}^d$ *then* $R_{y,X} = (R_{x^{(j)}}(y))$ *is a row vector and* $G_{y,X} = \left(G\left(y - x^{(j)}\right)\right)$ *is a row vector.*

Also $R_{X,y} = \left(R_y\left(x^{(i)}\right)\right)$ *is a column vector and* $G_{X,y} = \left(G\left(x^{(i)} - y\right)\right)$ *is a column vector.*

Using the notation of Definition 176 this theorem derives some of the relationships between the matrices $R_{X,X}$ and $G_{X,X}$.

Theorem 177 *Suppose* A *is any minimal unisolvent set and let the corresponding cardinal basis be* $\{l_i\}_{i=1}^M$. *Define the Riesz representer* R_x *using* A *and* $\{l_i\}_{i=1}^M$. *Then for any finite set* $X \subset \mathbb{R}^d$ *and* $y \in \mathbb{R}^d$

$$R_{y,X} = (2\pi)^{-\frac{d}{2}} \left(G_{y,X} - \widetilde{l}(y)^T G_{A,X} - G_{y,A} L_X^T + \widetilde{l}(y)^T G_{A,A} L_X^T\right) + \widetilde{l}(y)^T L_X^T.$$

The reproducing kernel matrix $R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right)$ *and the basis function matrix* $G_{X,X} = \left(G\left(x^{(i)} - x^{(j)}\right)\right)$ *are related by the formulas*

$$R_{X,X} = (2\pi)^{-\frac{d}{2}} \left(G_{X,X} - L_X G_{A,X} - G_{X,A} L_X^T + L_X G_{A,A} L_X^T\right) + L_X L_X^T, \tag{5.14}$$

and

$$R_{X,X} = (2\pi)^{-\frac{d}{2}} (I_N - L_{X;0}) G_{X,X} (I_N - L_{X;0})^T + L_X L_X^T, \tag{5.15}$$

where $L_X = \left(l_j(x^{(i)})\right)$ *and* $L_{X;0} = \begin{pmatrix} L_X & O_{N,N-M} \end{pmatrix}$.

Proof. From 4.35

$$(2\pi)^{\frac{d}{2}} R_x(y) = G(y-x) - \sum_{j=1}^M l_j(y) G\left(x^{(j)} - x\right) - \sum_{i=1}^M G\left(y - x^{(i)}\right) l_i(x) + \\ + \sum_{i,j=1}^M l_j(y) G\left(x^{(j)} - x^{(i)}\right) l_i(x) + (2\pi)^{\frac{d}{2}} \sum_{j=1}^M l_j(x) l_j(y),$$

or in the notation introduced in Definition 176

$$R_x(y) = (2\pi)^{-\frac{d}{2}} \left(G(y-x) - \widetilde{l}(y)^T G_{A,x} - G_{y,A} \widetilde{l}(x) + \widetilde{l}(y)^T G_{A,A} \widetilde{l}(x)\right) + \widetilde{l}(y)^T \widetilde{l}(x),$$

Now $R_{y,X}$ is the row vector $\left(R_{x^{(j)}}(y)\right)$ and $L_X = \left(l_j\left(x^{(i)}\right)\right)$ so

$$R_{y,X} = (2\pi)^{-\frac{d}{2}}\left(G_{y,X} - \widetilde{l}(y)^T G_{A,X} - G_{y,A}L_X^T + \widetilde{l}(y)^T G_{A,A}L_X^T\right) + \widetilde{l}(y)^T L_X^T,$$

and hence, since $L_{X;0} = (L_X \ O_{N,N-M})$,

$$\begin{aligned}
R_{X,X} &= \left(R_{x^{(i)},X}\right) \\
&= (2\pi)^{-\frac{d}{2}}\left(G_{x^{(i)},X} - \widetilde{l}\left(x^{(i)}\right)^T G_{A,X} - G_{x^{(i)},A}L_X^T + \widetilde{l}\left(x^{(i)}\right)^T G_{A,A}L_X^T\right) + \widetilde{l}\left(x^{(i)}\right)^T L_X^T \\
&= (2\pi)^{-\frac{d}{2}}\left(G_{X,X} - L_X^T G_{A,X} - G_{X,A}L_X^T + L_X G_{A,A}L_X^T\right) + L_X L_X^T \\
&= (2\pi)^{-\frac{d}{2}}\left(G_{X,X} - L_{X;0}G_{X,X} - G_{X,X}L_{X;0}^T + L_{X;0}G_{X,X}L_{X;0}^T\right) + L_{X;0}L_{X;0}^T \\
&= (2\pi)^{-\frac{d}{2}}\left(I_N - L_{X;0}\right)G_{X,X}\left(I_N - L_{X;0}\right)^T + L_{X;0}L_{X;0}^T.
\end{aligned}$$

■

5.5 Matrices, vectors and bases derived from the semi-Riesz representer r_x

The *semi-Riesz representer* $r_x = \mathcal{Q}R_x$ was introduced in Chapter 4 where $r_x(x)$ was used to estimate the convergence of the interpolant. In this document $r_x(x)$ will be used to estimate the convergence of the Exact smoother. Some basis properties of r_x were proved in Theorem 125 and we will need these to prove the following result:

Corollary 178 *Suppose $X = \left\{x^{(i)}\right\}_{i=1}^N$ is unisolvent and $A = \{a_i\}_{i=1}^M$ is a minimal unisolvent subset. If r_x is defined using this set then the functions $\left\{r_{x^{(i)}} : x^{(i)} \notin A\right\}$ are independent and form a basis for $\dot{W}_{G,X}$.*

Proof. Firstly, by part 2 Theorem 125, $r_{a_i} = 0$ for $i = 1, \ldots, M$. Next we prove independence. If $\sum\limits_{x^{(i)}\notin A} \beta_i r_{x^{(i)}} = 0$ then $\mathcal{Q} \sum\limits_{x^{(i)}\notin A} \beta_i R_{x^{(i)}} = 0$ so that $\sum\limits_{x^{(i)}\notin A} \beta_i R_{x^{(i)}} = \mathcal{P} \sum\limits_{x^{(i)}\notin A} \beta_i R_{x^{(i)}} \in P_\theta$. By part 6 of Summary 166, $\sum\limits_{x^{(i)}\notin A} \beta_i R_{x^{(i)}} \in \dot{W}_{G,X}$. But by part 3 of Summary 166, $W_{G,X} = \dot{W}_{G,X} \oplus P_\theta$ so all the β_i are zero. Also by part 3 of Summary 166, $\dot{W}_{G,X}$ has dimension $N-M$ and so the $N-M$ functions $\left\{r_{x^{(i)}} : x^{(i)} \notin A\right\}$ form a basis. ■

Definition 179 ***Matrices and vectors derived from*** r_x
If $X = \left\{x^{(k)}\right\} \subset \mathbb{R}^d$ and $y \in \mathbb{R}^d$ then $r_{X,X} = \left(r_{x^{(j)}}\left(x^{(i)}\right)\right)$, $r_{X,y} = \left(r_y\left(x^{(i)}\right)\right)$ is a column vector and $r_{y,X} = \left(r_{x^{(j)}}(y)\right)$ is a row vector.

Theorem 180 *Suppose $X = \left\{x^{(i)}\right\}_{i=1}^N$ is unisolvent and A is a minimal unisolvent subset with cardinal basis $\{l_i\}_{i=1}^M$. Then if $r_x = \mathcal{Q}R_x$ is defined using A we have:*

1. *$r_{X,y} = R_{X,y} - L_X\widetilde{l}(y)$.*
2. *$r_{X,X} = R_{X,X} - L_X L_X^T$.*
3. *The matrix r_{A^c,A^c} is positive definite, regular and Hermitian where $A^c = X\backslash A$.*

Proof. Parts 1 and 2 The definition of the cardinal unisolvency matrix is $L_X = \left(l_j\left(x^{(i)}\right)\right)$ and by part 6 Theorem 125, $R_y = r_y + \sum\limits_{j=1}^M l_j(y)\, l_j$. Parts 1 and 2 then follow easily from the definitions of $r_{X,y}$ and $r_{X,X}$.

Part 3 From Theorem 125, $r_{x^{(j)}}\left(x^{(i)}\right) = \left\langle r_{x^{(j)}}, r_{x^{(i)}}\right\rangle_{w,\theta}$ and $r_{x^{(j)}}(a_i) = 0$ for $a_i \in A$. Now let $(\cdot,\cdot)_{w,\theta}$ be the Light inner product 1.12: $(u,v)_{w,\theta} = \langle u,v\rangle_{w,\theta} + \sum\limits_{i=1}^M u(a_i)\overline{v(a_i)}$ constructed using A. It now follows that $r_{x^{(j)}}\left(x^{(i)}\right) = \left(r_{x^{(j)}}, r_{x^{(i)}}\right)_{w,\theta}$ for all i and j, and so the matrix $r_{A^c,A^c} = \left(r_{x^{(j)}}\left(x^{(i)}\right)\right)$ is a Gram matrix and the properties stated in this theorem are well-known properties of Gram matrices. ■

5.6 Matrix equations for the Exact smoother

In Corollary 172 it was shown that the unique solution of the Exact smoothing problem of the data $[X, y]$ lies in the finite dimensional space $W_{G,X}$ which has a unique representation given by 5.5 i.e.

$$\sum_{i=1}^{N} \alpha_i G\left(\cdot - x^{(i)}\right) + \sum_{j=1}^{M} \beta_j p_j, \quad \alpha_i, \beta_j \in \mathbb{C},$$

where $X = \left\{x^{(i)}\right\}_{i=1}^{N}$. The goal of this section is to derive a matrix equation for the coefficients α_i, β_j of the basis functions of the space $W_{G,X}$. We start by deriving a (singular) matrix equation for the values taken by the Exact smoother at its independent data points X. Now recall the notation f_X for $\widetilde{\mathcal{E}}_X f$ introduced in Definition 138.

Theorem 181 *Let $X = \left\{x^{(i)}\right\}_{i=1}^{N}$ be a unisolvent set of independent data with a minimally unisolvent subset A. Construct R_x using A and its cardinal basis $\{l_i\}_{i=1}^{M}$, and denote the Exact smoother of the data $[X, y]$ by s. Then*

$$\left(N\rho\left(I_N - L_{X;0}\right) + R_{X,X}\right) s_X = R_{X,X} y, \tag{5.16}$$

and

$$L_X^T \left(s_X - y\right) = 0, \tag{5.17}$$

where $L_X = \left(l_j\left(x^{(i)}\right)\right)$, $L_{X;0} = \begin{pmatrix} L_X & O_{N,N-M} \end{pmatrix}$ and $R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right)$ is the reproducing kernel matrix.

Proof. From parts 4 and 8 of Theorem 139 we have $\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* \beta = R_{X,X}\beta$ and $\widetilde{\mathcal{E}}_X \mathcal{P} s = L_X \widetilde{\mathcal{E}}_A s$. From part 1 of Corollary 172 we have $s = \mathcal{P}s - \frac{1}{N\rho}\widetilde{\mathcal{E}}_X^* \left(\widetilde{\mathcal{E}}_X s - y\right)$. Applying $\widetilde{\mathcal{E}}_X$ to the last equation to get

$$\begin{aligned} s_X = \widetilde{\mathcal{E}}_X \mathcal{P} s - \frac{1}{N\rho}\widetilde{\mathcal{E}}_X \widetilde{\mathcal{E}}_X^* \left(s_X - y\right) &= L_X \widetilde{\mathcal{E}}_A s - \frac{1}{N\rho} R_{X,X}\left(s_X - y\right) \\ &= L_{X;0} s_X - \frac{1}{N\rho} R_{X,X} s_X + \frac{1}{N\rho} R_{X,X} y, \end{aligned}$$

or on rearranging

$$\left(N\rho\left(I_N - L_{X;0}\right) + R_{X,X}\right) s_X = R_{X,X} y,$$

which is the desired matrix equation in $\widetilde{\mathcal{E}}_X s$. Finally 5.17 follows from part 3 Corollary 172 and equation 4.4 of part 3 Theorem 104. ■

Corollary 172 established that the Exact smoother has a unique solution in the finite dimensional space $W_{G,X}$. By Definition 130, $W_{G,X}$ is independent of the basis function G, the order of the points in X and the basis of P_θ used to define P_X. The next theorem derives the corresponding matrix equation for the coefficients of the basis functions. This matrix equation is deduced using the relationships between the reproducing kernel matrix $R_{X,X}$ and the basis function matrix $G_{X,X}$ derived in Theorem 177. However, to prove the next (well-known) result we need Lemma 143 from Chapter 4:

Lemma 182 *(Lemma 143)Let B be a complex-valued matrix and C be a real-valued matrix. Suppose the block matrix $\begin{pmatrix} B & C \\ C^T & O \end{pmatrix}$ is square and that for complex vectors z*

$$z^T B \overline{z} = 0 \text{ and } C^T z = 0 \text{ implies } z = 0. \tag{5.18}$$

1. *Then the equation*

$$\begin{pmatrix} B & C \\ C^T & O \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

   implies $u = 0$ and $v \in \operatorname{null} C$.

2. *If, in addition to 5.18 $\operatorname{null} C = \{0\}$, then the block matrix is regular.*

Theorem 183 *Suppose $[X, y]$ is the data for the Exact smoothing problem of Definition 167 where $X = \{x^{(i)}\}_{i=1}^{N}$ is θ-unisolvent and $y = \{y_i\}_{i=1}^{N}$. Then given a real-valued basis $\{p_j\}_{j=1}^{M}$ for P_θ and a basis function G of order θ, the Exact smoothing problem has a unique solution $s \in W_{G,X}$ of the form*

$$s(x) = \sum_{i=1}^{N} v_i G\left(x - x^{(i)}\right) + \sum_{j=1}^{M} \beta_j p_j(x), \tag{5.19}$$

where the coefficients $v = (v_i) \in \mathbb{C}^N$ and $\beta = \left(\beta_j\right) \in \mathbb{C}^N$ satisfy the ***Exact smoothing matrix equation***

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X} & P_X \\ P_X^T & O_M \end{pmatrix} \begin{pmatrix} v \\ \beta \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}. \tag{5.20}$$

Here $P_X = \left(p_j\left(x^{(i)}\right)\right)$ is a unisolvency matrix and $G_{X,X} = \left(G\left(x^{(i)} - x^{(j)}\right)\right)$ is the basis function matrix.

The matrix equation 5.20 is independent of the ordering of the data $[X, y]$.

Finally, the Exact smoothing matrix is regular and positive definite but in general it is not Hermitian. However, there always exists a basis function such that the Exact smoothing matrix is Hermitian.

Proof. Step 1 Since $X = \{x^{(i)}\}_{i=1}^{N}$ is θ-unisolvent there exists a minimal unisolvent subset, say X_1. Assume that $X_1 = \{x^{(i)}\}_{i=1}^{M}$. Now construct $\mathcal{P}$, $\mathcal{Q}$, the Light norm and the Riesz representer R_x using X_1 and its cardinal basis $\{l_i\}_{i=1}^{M}$ for P_θ.

Initially we will derive the Exact smoother matrix equation using the cardinal basis associated with X_1 so that $P_X = L_X = \left(l_j\left(x^{(i)}\right)\right)$. Consequently, by the interpolation result Theorem 144 of Chapter 4, for given dependent data $s_X = \left(s\left(x^{(i)}\right)\right)_{i=1}^{N}$ there exist unique vectors v and γ such that

$$\begin{pmatrix} G_{X,X} & L_X \\ L_X^T & O_M \end{pmatrix} \begin{pmatrix} v \\ \gamma \end{pmatrix} = \begin{pmatrix} s_X \\ 0 \end{pmatrix},$$

This matrix equation is equivalent to the two equations

$$G_{X,X} v + L_X \gamma = s_X \tag{5.21}$$

$$L_X^T v = 0. \tag{5.22}$$

Multiplying equation 5.21 by $N\rho R_{X,X}^{-1}\left(I_N - L_{X;0}\right) + I_N$ and using equation 5.16 gives

$$\left(N\rho R_{X,X}^{-1}\left(I_N - L_{X;0}\right) + I_N\right)\left(G_{X,X} v + L_X \gamma\right) = \left(N\rho R_{X,X}^{-1}\left(I_N - L_{X;0}\right) + I_N\right) s_X = y.$$

From part 3 Theorem 105, $\left(I_N - L_{X;0}\right) L_X = O$ so the last equation simplifies to

$$N\rho R_{X,X}^{-1}\left(I_N - L_{X;0}\right) G_{X,X} v + G_{X,X} v + L_X \gamma = y. \tag{5.23}$$

We now require an equation that expresses $R_{X,X}$ in terms of $G_{X,X}$. To this end we use equation 5.15, namely

$$R_{X,X} = (2\pi)^{-\frac{d}{2}}\left(I_N - L_{X;0}\right) G_{X,X}\left(I_N - L_{X;0}\right)^T + L_X L_X^T.$$

Multiplying this equation on the left by $R_{X,X}^{-1}$ and by v on the right, and noting that $L_X^T v = 0$, yields

$$\begin{aligned} v &= (2\pi)^{-\frac{d}{2}} R_{X,X}^{-1}\left(I_N - L_{X;0}\right) G_{X,X}\left(I_N - L_{X;0}\right)^T v + R_{X,X}^{-1} L_X L_X^T v \\ &= (2\pi)^{-\frac{d}{2}} R_{X,X}^{-1}\left(I_N - L_{X;0}\right) G_{X,X} v, \end{aligned}$$

and thus 5.23 reduces to

$$\left((2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X}\right) v + L_X \gamma = y, \tag{5.24}$$

constrained by equation 5.22 i.e. $L_X^T v = 0$.

Since $P_X = \left(p_j\left(x_i\right)\right)$ and $L_X = \left(l_j\left(x_i\right)\right)$, where $\{p_j\}$ and $\{l_j\}$ are both real-valued bases for P_θ, it follows from part 3 Theorem 104 that there exists a regular matrix C such that $L_X = P_X C$. Substituting for L_X in 5.24 and 5.22 and then setting $\beta = C\gamma$ yields the Exact smoothing matrix equation 5.20.

Step 2 We now show that re-ordering the data does not change the equation for the smoother. Re-ordering a vector involves a permutation π. Re-ordering a column vector involves left-multiplication by the permutation matrix Π and re-ordering a row vector involves right-multiplication by the transpose of the permutation matrix. Thus the new matrix equation is

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho I_N + G_{\pi(X),\pi(X)} & P_{\pi(X)} \\ P_{\pi(X)}^T & O_M \end{pmatrix} \begin{pmatrix} v' \\ \beta' \end{pmatrix} = \begin{pmatrix} \Pi y \\ 0 \end{pmatrix}.$$

Thus $G_{\pi(X),\pi(X)} = \Pi G_{X,X} \Pi^T$, $P_{\pi(X)} = \Pi P_X$ and $\pi(y) = \Pi y$, or since $\Pi\Pi^T = I_N$. Under the last set of transformations the Exact smoothing matrix equation becomes

$$\left((2\pi)^{\frac{d}{2}} N\rho\Pi\Pi^T + \Pi G_{X,X}\Pi^T\right) v' + \Pi P_X \beta' = \Pi y, \quad P_X^T \Pi^T v' = 0,$$

or

$$\left((2\pi)^{\frac{d}{2}} N\rho + G_{X,X}\right) \Pi^T v' + P_X \beta' = y, \quad P_X^T \Pi^T v' = 0, \tag{5.25}$$

and so the matrix equation is unchanged by re-ordering the data.

Step 3 The next step is to show that the Exact smoothing matrix is regular. To do this we use Lemma 182 with $B = (2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X}$, $C = P_X$, and then show that $z^T \left((2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X}\right) \overline{z} = 0$ and $P_X^T z = 0$ implies $z = 0$. Now

$$\begin{aligned} z^T \left((2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X}\right) \overline{z} &= (2\pi)^{\frac{d}{2}} z^T N\rho I_N \overline{z} + z^T G_{X,X} \overline{z} \\ &= (2\pi)^{\frac{d}{2}} N\rho \left|z\right|^2 + z^T G_{X,X} \overline{z}. \end{aligned}$$

But by part 2 Theorem 166, $G_{X,X}$ is conditionally positive definite on null P_X^T i.e. $z \in$ null P_X^T implies that $z^T G_{X,X} \overline{z} > 0$ except when $z = 0$. Thus $z \in$ null P_X^T and $z^T \left((2\pi)^{\frac{d}{2}} N\rho I_N + G_{X,X}\right) \overline{z} = 0$ implies $z = 0$ and the Approximate smoother matrix is regular.

Finally, part 2 Theorem 92 from Chapter 3 allows the basis function G to be chosen so that $\overline{G(x)} = G(-x)$ and this implies $G_{X,X}$ is Hermitian. ■

Remark 184

1. *When $\rho = 0$ the matrix equation for the Exact smoother becomes the matrix equation for the minimal norm interpolant - see Theorem 144.*
2. *The matrix 5.20 is $N \times N$ i.e. its size depends on the number of data points, so this algorithm is **not scalable** i.e. the time of execution is not linearly dependent on the number of data points. The **Approximate smoother**, which overcomes this problem, will be derived in Chapter 6.*

The basis function form of the Exact smoother matrix equation can also be derived using Lagrange multipliers. The next result shows that the algebra can be significantly simplified by assuming the basis function is real valued.

Corollary 185 *If the dependent data y is real-valued and the basis function is real-valued, then the Exact smoother is real-valued. Also, the smoother lies in the subspace of $W_{G,X}$ defined using the real scalars instead of the complex scalars.*

Proof. By Theorem 177, if the basis function G is real-valued then the reproducing kernel matrix $R_{X,X}$ is real-valued. Hence, since the cardinal unisolvency matrix L_X is real-valued, equations 5.19 and 5.20 imply that the smoother is real-valued. ■

5.7 Convergence to the data function - smoother error

In this section we will prove that in the sense of Corollary 190 below the Exact smoother converges uniformly pointwise to its data function on a bounded set. The rates of convergence are shown to be the same as those obtained for the minimal seminorm interpolant.

5.7.1 Convergence to the data function

In Section 5.3 the Exact smoother problem of Definition 168 was studied using the Hilbert space $V = X_w^\theta \otimes \mathbb{C}^N$ endowed with the inner product $\left((f,\widetilde{\alpha}), \left(g,\widetilde{\beta}\right)\right)_V = \rho \langle f,g\rangle_{w,\theta} + \frac{1}{N}\left(\widetilde{\alpha},\widetilde{\beta}\right)_{\mathbb{C}^N}$, and the operator $\mathcal{L}_X : X_w^\theta \to V$ defined by $\mathcal{L}_X f = \left(f, \widetilde{\mathcal{E}}_X f\right)$ for a set of N independent data points X.

In Theorem 170 it was proven that $\|\mathcal{L}_X f\|_V$ is an equivalent norm to $\|f\|_{w,\theta}$ which implies that X_w^θ is also a reproducing kernel Hilbert space under the norm $\|\mathcal{L}_X f\|_V$. The space V induces on X_w^θ the inner product

$$(f,g)_{V,w,\theta} = (\mathcal{L}_X f, \mathcal{L}_X g)_V\,, \quad f,\ g \in X_w^\theta.$$

Under this inner product X_w^θ is a reproducing kernel Hilbert space with a unique reproducing kernel function, and consequently there is a unique Riesz representer of the functional $f \to f(x)$ which we denote by $\mathfrak{R}_{V,x}$ i.e.

$$f(x) = (f, \mathfrak{R}_{V,x})_{V,w,\theta}\,, \quad f \in X_w^\theta,\ x \in \mathbb{R}^d. \tag{5.26}$$

The equations

$$(f,R_x)_{w,\theta} = f(x) = (f,\mathfrak{R}_{V,x})_{V,w,\theta} = (\mathcal{L}_X f, \mathcal{L}_X \mathfrak{R}_{V,x})_V = (f, \mathcal{L}_X^* \mathcal{L}_X \mathfrak{R}_{V,x})_{w,\theta}\,, \tag{5.27}$$

imply that $\mathcal{L}_X^* \mathcal{L}_X \mathfrak{R}_{V,x} = R_x$ and by Theorem 175 the operator $\mathcal{L}_X^* \mathcal{L}_X : X_w^\theta \to X_w^\theta$ is a homeomorphism so we have

$$\mathfrak{R}_{V,x} = \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} R_x. \tag{5.28}$$

The equations 5.27 also imply there exists a unique $R_{V,x} \in V$ such that

$$f(x) = (\mathcal{L}_X f, R_{V,x})_V\,, \quad R_{V,x} = \mathcal{L}_X \mathfrak{R}_{V,x}, \quad R_x = \mathcal{L}_X^* R_{V,x}. \tag{5.29}$$

As mentioned above we are interested in estimating the pointwise error of the Exact smoother s_e with respect to its data function f_d, where the independent data is X and the dependent data is $y = \widetilde{\mathcal{E}}_X f_d$. By 5.29 the error is $s_e(x) - f_d(x) = (\mathcal{L}_X(s_e - f_d), R_{V,x})_V$ and

$$|s_e(x) - f_d(x)| = \left|(\mathcal{L}_X(s_e - f_d), R_{V,x})_V\right| \le \|\mathcal{L}_X(s_e - f_d)\|_V \|R_{V,x}\|_V\,.$$

From part 3 of Theorem 171

$$\|\mathcal{L}_X s_e - \varsigma\|_V^2 + \|\mathcal{L}_X(s_e - f)\|_V^2 = \|\mathcal{L}_X f - \varsigma\|_V^2\,, \quad f \in X_w^\theta,$$

where $\zeta = (0,y) = \left(0, \widetilde{\mathcal{E}}_X f_d\right)$. Thus when $f = f_d$

$$\|\mathcal{L}_X(s_e - f_d)\|_V \le \left\|\mathcal{L}_X f_d - \left(0, \widetilde{\mathcal{E}}_X f_d\right)\right\|_V = |f_d|_{w,\theta}\sqrt{\rho},$$

and

$$|s_e(x) - f_d(x)| \le |f_d|_{w,\theta}\sqrt{\rho}\,\|R_{V,x}\|_V\,.$$

Finally, using 5.27 we have

$$\|R_{V,x}\|_V^2 = (R_{V,x}, R_{V,x})_V = (\mathcal{L}_X \mathfrak{R}_{V,x}, \mathcal{L}_X \mathfrak{R}_{V,x})_V = \mathfrak{R}_{V,x}(x)\,.$$

The above analysis is summarized as:

Theorem 186 *Suppose s_e is the Exact smoother generated by the independent data X and the data function $f_d \in X_w^\theta$. Then*

$$|s_e(x) - f_d(x)| \le |f_d|_{w,\theta}\sqrt{\rho}\,\|R_{V,x}\|_V\,, \quad x \in \mathbb{R}^N, \tag{5.30}$$

where $\rho > 0$ is the smoothing coefficient, $R_{V,x} \in V$ is given by 5.28 and 5.29 and satisfies

$$\|R_{V,x}\|_V^2 = \mathfrak{R}_{V,x}(x)\,. \tag{5.31}$$

It is clear from Definition 167 of the Exact smoothing problem that the Exact smoother is independent of the order of the points in the unisolvent independent data X. This allows the convenient definition of a special minimal unisolvent set $X_1 \subset X$: we assume that $X_1 = \left\{x^{(i)}\right\}_{i=1}^{M}$ is a minimal unisolvent subset and let $X_2 = \left\{x^{(i)}\right\}_{i=M+1}^{N}$. The set X_1 is then used to construct the Riesz representer R_x, the semi-Riesz representer $r_x = \mathcal{Q}R_x$, the Lagrangian operators $\mathcal{P}$, $\mathcal{Q}$, the basis function spaces $W_{G,X}$ and $\dot{W}_{G,X}$, and the Light norm $(\cdot,\cdot)_{w,\theta}$. This section will also use the 'prime-double-prime' notation to denote components related to X_1 and X_2 e.g. $\alpha = (\alpha', \alpha'')$ with $\alpha' \in \mathbb{C}^M$ and $\alpha'' \in \mathbb{C}^{N-M}$.

The next step is to calculate $\|R_{V,x}\|_V$ using 1.62. From 5.28, $\mathcal{L}_X^* \mathcal{L}_X \mathfrak{R}_{V,x} = R_x$ and from part 4 Theorem 170, $\mathcal{L}_X^* \mathcal{L}_X = \rho\mathcal{Q} + \frac{1}{N}\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X$ so that

$$\rho\mathcal{Q}\mathfrak{R}_{V,x} + \frac{1}{N}\widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X \mathfrak{R}_{V,x} = R_x. \tag{5.32}$$

We will solve equation 5.32 for $\mathfrak{R}_{V,x}$ by solving the equivalent system

$$\begin{aligned} \mathcal{Q}v_x + \widetilde{\mathcal{E}}_X^* \alpha_x &= R_x, && (5.33) \\ \frac{1}{N\rho}\widetilde{\mathcal{E}}_X v_x &= \alpha_x, && (5.34) \end{aligned}$$

for $\alpha_x \in \mathbb{C}^N$ and $v_x \in X_w^\theta$, so that

$$\rho\mathfrak{R}_{V,x} = v_x,\ R_{V,x} = \mathcal{L}_X \mathfrak{R}_{V,x} = \left(\frac{1}{\rho}v_x, N\alpha_x\right). \tag{5.35}$$

We first apply the operator $\mathcal{P}$ to 5.33: by part 7 of Theorem 139, $\mathcal{P}\widetilde{\mathcal{E}}_X^* \alpha_x = \alpha_x^T L_X \widetilde{l}$ and applying $\mathcal{P}$ to 4.36 gives $\mathcal{P}R_x = \widetilde{l}\,(x)^T\, \widetilde{l}$ so

$$\mathcal{P}\left(\mathcal{Q}v_x + \widetilde{\mathcal{E}}_X^* \alpha_x\right) = \alpha_x^T L_X \widetilde{l} = \widetilde{l}\,(x)^T\, \widetilde{l},$$

which implies the equivalent equations

$$L_X^T \alpha_x = \widetilde{l}\,(x)\,, \quad \alpha_x' = \widetilde{l}\,(x) - L_{X_2}^T \alpha_x''. \tag{5.36}$$

Next apply the operator $\mathcal{Q}$ to 5.33: since $\mathcal{Q}R_x = r_x$

$$\mathcal{Q}\left(v_x + \widetilde{\mathcal{E}}_X^* \alpha_x\right) = \mathcal{Q}v_x + \mathcal{Q}\sum_{k=1}^{N} (\alpha_x)_k\, R_{x^{(k)}} = \mathcal{Q}v_x + \sum_{k=M+1}^{N} (\alpha_x)_k\, r_{x^{(k)}} = r_x,$$

and thus

$$\mathcal{Q}v_x = r_x - \sum_{k=M+1}^{N} (\alpha_x)_k\, r_{x^{(k)}}. \tag{5.37}$$

We now left-compose $\widetilde{\mathcal{E}}_{X_2}$ with 5.37. From part 8 Theorem 139, $\widetilde{\mathcal{E}}_{X_2}\mathcal{P}f = L_X \widetilde{\mathcal{E}}_{X_1} f$ so that applying $\widetilde{\mathcal{E}}_{X_2}$ to the left of 5.37 and then using 1.30 gives

$$\widetilde{\mathcal{E}}_{X_2}\mathcal{Q}v_x = \widetilde{\mathcal{E}}_{X_2}\left(v_x - \mathcal{P}v_x\right) = \widetilde{\mathcal{E}}_{X_2} v_x - L_{X_2}\widetilde{\mathcal{E}}_{X_1} v_x = N\rho\left(\alpha_x'' - L_{X_2}\alpha_x'\right).$$

Next left-compose $\widetilde{\mathcal{E}}_{X_2}$ with 5.37 and use the notation described in Definition 179 to obtain

$$\widetilde{\mathcal{E}}_{X_2}\mathcal{Q}v_x = \widetilde{\mathcal{E}}_{X_2}\left(r_x - \sum_{k=M+1}^{N} (\alpha_x)_k\, r_{x^{(k)}}\right) = r_{X_2,x} - r_{X_2,X_2}\alpha_x''.$$

The last two sequences of equations yield

$$\left(N\rho I + r_{X_2,X_2}\right)\alpha_x'' = N\rho L_{X_2}\alpha_x' + r_{X_2,x},$$

so that substituting for α_x' in the last equation using 5.36 and then rearranging gives

$$\left(N\rho\left(I + L_{X_2}L_{X_2}^T\right) + r_{X_2,X_2}\right)\alpha_x'' = N\rho L_{X_2}\widetilde{l}\,(x) + r_{X_2,x}. \tag{5.38}$$

We know from Theorem 180 that r_{X_2,X_2} is positive definite and since the cardinal unisolvency matrix L_{X_2} is real-valued it follows that $N\rho\left(I+L_{X_2}L_{X_2}^T\right)+r_{X_2,X_2}$ is also positive definite and hence regular. Thus 5.38 and 5.36 define α_x uniquely and which means 5.33 defines v up to a polynomial of order θ. However 5.34 actually supplies the information which defines v_x uniquely. Indeed, from 5.34, $\widetilde{\mathcal{E}}_{X_1}v_x = N\rho\alpha_x'$ so that

$$\begin{aligned}\mathcal{P}v_x\left(y\right) = N\rho\ \left(\alpha_x'\right)^T\widetilde{l}\left(y\right) = N\rho\left(\widetilde{l}\left(x\right)-L_{X_2}^T\alpha_x''\right)^T\widetilde{l}\left(y\right) &= N\rho\left(\widetilde{l}\left(x\right)^T-\left(\alpha_x''\right)^T L_{X_2}\right)\widetilde{l}\left(y\right)\\ &= N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-N\rho\left(\alpha_x''\right)^T L_{X_2}\widetilde{l}\left(y\right),\end{aligned}$$

and as a consequence of 5.37

$$\begin{aligned}v_x\left(y\right) &= \mathcal{P}v_x\left(y\right)+\mathcal{Q}v_x\left(y\right)\\ &= N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-N\rho\left(\alpha_x''\right)^T L_{X_2}\widetilde{l}\left(y\right)+\mathcal{Q}v_x\left(y\right)\\ &= N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-N\rho\left(\alpha_x''\right)^T L_{X_2}\widetilde{l}\left(y\right)+r_x\left(y\right)-\sum_{k=M+1}^{N}\left(\alpha_x\right)_k r_{x^{(k)}}\left(y\right)\\ &= r_x\left(y\right)+N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-N\rho\left(\alpha_x''\right)^T L_{X_2}\widetilde{l}\left(y\right)-\left(\alpha_x''\right)^T\overline{r_{X_2,y}}\\ &= r_x\left(y\right)+N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-\left(\alpha_x''\right)^T\left(N\rho L_{X_2}\widetilde{l}\left(y\right)+\overline{r_{X_2,y}}\right)\\ &= r_x\left(y\right)+N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-\left(\alpha_x''\right)^T\left(N\rho\ \left(I+L_{X_2}L_{X_2}^T\right)+\overline{r_{X_2,X_2}}\right)\overline{\alpha_y''} \qquad (5.39)\\ &= r_x\left(y\right)+N\rho\widetilde{l}\left(x\right)^T\widetilde{l}\left(y\right)-\\ &\quad-\left(N\rho L_{X_2}\widetilde{l}\left(x\right)+r_{X_2,x}\right)^T\left(N\rho\ \left(I+L_{X_2}L_{X_2}^T\right)+\overline{r_{X_2,X_2}}\right)^{-1}\left(N\rho L_{X_2}\widetilde{l}\left(y\right)+\overline{r_{X_2,y}}\right), \qquad (5.40)\end{aligned}$$

where the last two equations were derived using 5.38. Since $N\rho\left(I+L_{X_2}L_{X_2}^T\right)+r_{X_2,X_2}$ is positive definite, its complex conjugate inverse is positive definite and when $y=x$ equations 5.31, 5.39 and 5.35 imply

$$\rho\left\|R_{V,x}\right\|_V^2=\rho\mathfrak{R}_{V,x}\left(x\right)=v_x\left(x\right)\leq r_x\left(x\right)+N\rho\left|\widetilde{l}\left(x\right)\right|^2,$$

The next theorem summarizes these results:

Theorem 187 *Suppose s_e is the Exact smoother generated by the independent data X and the data function $f_d\in X_w^\theta$. Then*

$$\rho\left\|R_{V,x}\right\|_V^2\leq r_x\left(x\right)+N\rho\left|\widetilde{l}\left(x\right)\right|^2, \qquad (5.41)$$

and

$$\left|s_e\left(x\right)-f_d\left(x\right)\right|\leq\left|f_d\right|_{w,\theta}\sqrt{r_x\left(x\right)+N\rho\left|\widetilde{l}\left(x\right)\right|^2},\quad x\in\mathbb{R}^N.$$

To study the convergence of the Exact smoother we will also need Lemma 148 which supplies some elementary results from the theory of Lagrange interpolation and was used in Chapter 4 to derive orders of convergence for the interpolant. These results are stated without proof. This lemma has been created from Lemma 3.2, Lemma 3.5 and the first two paragraphs of the proof of Theorem 3.6 of Light and Wayne [10]. The results of this lemma do not involve any reference to weight or basis functions or functions in X_w^θ, but consider the properties of the set which contains the independent data points and the order of the unisolvency used for the interpolation. Thus we have separated the part of the proof that involves weight functions from the part that uses the detailed theory of Lagrange interpolation operators.

Lemma 188 *(Copy of Lemma 148) Suppose that:*

1. *Ω is a bounded, open, connected subset of $\mathbb{R}^d$ having the cone property.*
2. *X is a unisolvent subset of Ω of order θ.*
3. *$\{l_j\}_{j=1}^M$ is the cardinal basis of P_θ with respect to a minimal unisolvent set of Ω.*

   Then by using Lagrange interpolation techniques it can be shown there exists a constant $K_{\Omega,\theta}'>0$ such that

$$\left|\widetilde{l}\left(x\right)\right|_1=\sum_{j=1}^{M}\left|l_j\left(x\right)\right|\leq K_{\Omega,\theta}',\quad x\in\overline{\Omega}, \qquad (5.42)$$

and all minimal unisolvent subsets of Ω. *Now define*

$$h_X = \sup_{\omega \in \Omega} \operatorname{dist}(\omega, X),$$

and fix $x \in X$. *Again using Lagrange interpolation techniques it can be shown there are constants* $c_{\Omega,\theta}, h_{\Omega,\theta} > 0$ *such that when* $h_X < h_{\Omega,\theta}$ *there exists a minimal unisolvent set* $A \subset X$ *satisfying*

$$\operatorname{diam} A_x \leq c_{\Omega,\theta} h_X, \tag{5.43}$$

where $A_x = A \cup \{x\}$.

We will now prove our first Exact smoother error estimate Theorem 189 which implies an order of convergence η. In Theorem 150 conditions on the basis function G were supplied in order that the semi-Riesz representer $r_x(y) = \mathcal{Q}_y R_x(y)$ satisfy a condition of the form 5.44 required below by Theorem 189. Then in Theorem 152 an estimate for the interpolant error was derived which is just the Exact smoother error estimate 5.45 with $\rho = 0$.

Theorem 189 *Let* w *be a weight function with properties W2 and W3 for order* θ *and parameter* κ *and set* $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. *Suppose the notation and assumptions of Lemma 188 hold. Assume there exist constants* $c_G, r_G > 0$, *independent of* A *and* $x \in \Omega$ *such that*

$$\sqrt{r_x(x)} \leq \left(1 + \left|\widetilde{l}(x)\right|_1\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^\eta, \quad \operatorname{diam} A_x < r_G,\ x \in \Omega, \tag{5.44}$$

where $A_x = A \cup \{x\}$.

Then there exist constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta} > 0$ *such that*

$$|s_e(x) - f_d(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^\eta + \sqrt{N_X \rho}\right), \quad x \in \overline{\Omega}, \tag{5.45}$$

when $h_X \leq \min\{h_{\Omega,\theta}, r_G\}$. *Here* $N_X = |X|$ *and the constants* $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta}$ *only depend on* Ω, θ, κ *and* d.

Proof. From 5.41 of Theorem 187

$$\sqrt{\rho} \left\|R_{V,x}\right\|_V \leq \sqrt{r_x(x)} + \left|\widetilde{l}(x)\right|_1 \sqrt{N_X \rho}, \quad x \in \mathbb{R}^d. \tag{5.46}$$

Fix $x \in \nleq$ and let A be any minimal unisolvent subset of X. Define r_x using A so that by 5.44 and Lemma 188 we have

$$\sqrt{r_x(x)} \leq \sqrt{c_G} \left(1 + \left|\widetilde{l}(x)\right|_1\right) \left(\operatorname{diam} A_x\right)^\eta \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta + \delta_G}, \tag{5.47}$$

when $\operatorname{diam} A_x \leq r_G$. But $h_X \leq h_{\Omega,\theta}$ so that by Lemma 188 there exists A such that $\operatorname{diam} A_x \leq c_{\Omega,\theta} h_X$ and 5.47 implies

$$\sqrt{r_x(x)} \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^\eta.$$

Applying this estimate and the inequality 5.42 for $\left|\widetilde{l}(x)\right|_1$ to the inequality 5.46 for $R_{V,x}$ we get

$$\sqrt{\rho} \left\|R_{V,x}\right\|_V \leq \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^\eta + \sqrt{N_X \rho}\right), \quad x \in \Omega,$$

and the smoother error estimate 5.30 now becomes

$$\begin{aligned} |s_e(x) - f_d(x)| &\leq |f_d|_{w,\theta} \sqrt{\rho} \left\|R_{V,x}\right\|_V \\ &\leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^\eta + \sqrt{N_X \rho}\right), \quad x \in \Omega. \end{aligned}$$

Since s_e and f_d are continuous on $\mathbb{R}^d$ the last inequality actually holds for $x \in \overline{\Omega}$ and 5.45 is true. ■

The next result shows that in a certain sense the Exact smoother converges to its data function.

Corollary 190 *The Exact smoother converges to its data function* f_d *in the sense that given* $\varepsilon > 0$ *there exists a positive integer* $K(f_d; \varepsilon)$, *a nested sequence of independent data sets* $X^{(k)} \subset X^{(k+1)} \subset \Omega$ *and a sequence of smoothing parameters* $\rho_k > 0$ *such that the corresponding sequence of Exact smoothers* $s_e^{(k)}$ *satisfies*

$$\left|s_e^{(k)}(x) - f_d(x)\right| \leq \varepsilon, \quad x \in \overline{\Omega}, \tag{5.48}$$

when $k \geq K(f_d; \varepsilon)$.

Proof. From Theorem 151 there exists a nested sequence of independent data sets $X^{(k)}$ such that $X^{(k)} \subset X^{(k+1)} \subset \Omega$ and $h_{X^{(k)}} \to 0$ as $k \to \infty$. To derive the bound 5.48 we start with inequality 5.45 so that

$$\left| s_e^{(k)}(x) - f_d(x) \right| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_{X^{(k)}}\right)^{\eta_G} + \sqrt{N_{X^{(k)}} \rho_k} \right), \quad x \in \overline{\Omega},$$

when $h_{X^{(k)}} < \min\{h_{\Omega,\theta}, r_G\}$. Now observe that for some positive integer $K_0(f_d)$, $h_{X^{(k)}} < h_{\Omega,\theta}$ when $k \geq K_0(f_d)$. Next, choosing ρ_k such that $\sqrt{c_G}\left(c_{\Omega,\theta} h_{X^{(k)}}\right)^{\eta_G} = \sqrt{N_{X^{(k)}} \rho_k}$ implies

$$\left| s_e^{(k)}(x) - f_d(x) \right| \leq 2 |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta} h_{X^{(k)}}\right)^{\eta_G}, x \in \overline{\Omega},$$

and as $h_{X^{(k)}} \to 0$, $s_e \to f_d$ uniformly on $\overline{\Omega}$. The statement of this corollary now follows directly. ■

5.8 Improved error estimates

In this section we start by deriving an estimate which has a slightly improved order of convergence: for the example of the shifted thin-plate spline this is 1/2. Then a double rate of convergence will be demonstrated for data functions that are linear combinations of Riesz representers. As a function of the smoothing parameter ρ these convergence estimates are unbounded and we end this chapter by deriving some error estimates bounded in ρ.

5.8.1 A slightly increased order of convergence

Our slightly improved Exact smoother error estimate will be Theorem 191. In Theorem 156 conditions on the basis function G were supplied in order that the semi-Riesz representer $r_x(y) = \mathcal{Q}_y R_x(y)$ satisfy a condition of the form 5.49 (below) which allows for an improved order of convergence of $\eta + \delta_G$. Then in Corollary 157 an estimate for the interpolant error was derived which is just the improved Exact smoother error estimate 5.50 (below) with $\rho = 0$.

Theorem 191 *Let w be a weight function with properties W2 and W3 for order θ and parameter κ and set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Suppose the notation and assumptions of Lemma 188 hold. Assume there exist constants $c_G, r_G > 0$ and $\delta_G \geq 0$, independent of A and $x \in \Omega$, such that*

$$\sqrt{r_x(x)} \leq \left(1 + \left|\widetilde{l}(x)\right|_1\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta + \delta_G}, \quad \operatorname{diam} A_x < r_G,\ x \in \Omega, \tag{5.49}$$

where $A_x = A \cup \{x\}$.

Then there exist constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta} > 0$ such that

$$|s_e(x) - f_d(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta + \delta_G} + \sqrt{N_X \rho} \right), \quad x \in \overline{\Omega}, \tag{5.50}$$

when $h_X \leq \min\{h_{\Omega,\theta}, r_G\}$. Here $N_X = |X|$ and the constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta}$ only depend on Ω, θ, κ and d.

Proof. The will follow closely that of Theorem 189. From 5.41 of Theorem 187

$$\sqrt{\rho} \left\| R_{V,x} \right\|_V \leq \sqrt{r_x(x)} + \left|\widetilde{l}(x)\right|_1 \sqrt{N_X \rho}, \quad x \in \mathbb{R}^d. \tag{5.51}$$

Fix $x \in \not\leq$ and let A be any minimal unisolvent subset of X. Define r_x using A so that by 5.49 and Lemma 188 we have

$$\sqrt{r_x(x)} \leq \sqrt{c_G} \left(1 + \left|\widetilde{l}(x)\right|_1\right) \left(\operatorname{diam} A_x\right)^{\eta + \delta_G} \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta + \delta_G}, \tag{5.52}$$

when $\operatorname{diam} A_x \leq r_G$. But $h_X \leq h_{\Omega,\theta}$ so that by Lemma 188 there exists A such that $\operatorname{diam} A_x \leq c_{\Omega,\theta} h_X$ and 5.52 implies

$$\sqrt{r_x(x)} \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta + \delta_G}. \tag{5.53}$$

Applying this estimate and the inequality 5.42 for $\left|\widetilde{l}(x)\right|_1$ to the inequality 5.51 for $R_{V,x}$ we get

$$\sqrt{\rho} \left\| R_{V,x} \right\|_V \leq \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta + \delta_G} + \sqrt{N_X \rho} \right), \quad x \in \Omega, \tag{5.54}$$

and the smoother error estimate 5.30 now becomes

$$\begin{aligned}|s_e(x)-f_d(x)| &\leq |f_d|_{w,\theta}\sqrt{\rho}\,\|R_{V,x}\|_V \\ &\leq |f_d|_{w,\theta}\left(1+K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\,(c_{\Omega,\theta}h_X)^{\eta+\delta_G}+\sqrt{N_X\rho}\right), \quad x\in\Omega.\end{aligned}$$

Since s_e and f_d are continuous on $\mathbb{R}^d$ the last inequality actually holds for $x\in\overline{\Omega}$. ■

In the examples below, to maximize the rate of convergence, we define $\eta=\max\limits_{0\leq\kappa<s}\min\left\{\theta,\frac{1}{2}\lfloor 2\kappa\rfloor\right\}$.

Example 192 ***Thin plate spline basis functions*** *By Theorem 71 the thin-plate spline weight functions are given by*

$$w(\xi)=\frac{1}{e(s)}|\xi|^{-2\theta+2s+d}, \quad \xi\in\mathbb{R}^d, \tag{5.55}$$

and have properties W2.1 and W3.2 with positive integer order θ and non-negative κ iff $\kappa<s<\theta$.

Now $\eta=\max\limits_{0\leq\kappa<s}\min\left\{\theta,\frac{1}{2}\lfloor 2\kappa\rfloor\right\}=\frac{1}{2}\lfloor 2\kappa\rfloor$ and from the conclusion of Example 1 of Section 4.11 Chapter 4, equation 5.49 is satisfied for $r_G=1$ and $\delta_G=\frac{1}{2}(2s-\lfloor 2s\rfloor)$ when $s>0$, $s\neq 1,2,3,\ldots$ and for $r_G=1$ and any $0\leq\delta_G<\frac{1}{2}$ when $s=1,2,3,\ldots$.

Example 193 ***Shifted thin plate spline basis functions*** *By Theorem 73 the shifted thin-plate spline weight functions are given by*

$$w(\xi)=\frac{1}{\widetilde{e}(s)\,\widetilde{K}_{s+d/2}(a|\xi|)}|\xi|^{-2\theta+2s+d}, \quad s>-d/2,$$

and have properties W2.1 and W3.2 for θ and all $\kappa\geq 0$ iff $-d/2<s<\theta$.

Now $\eta=\max\limits_{\kappa\geq 0}\min\left\{\theta,\frac{1}{2}\lfloor 2\kappa\rfloor\right\}=\theta$ and from the conclusion of Example 2 of Section 4.11 Chapter 4, 5.49 is satisfied for arbitrary $r_G>0$ and $\delta_G=\frac{1}{2}$.

Remark 194

1. $\boxed{\textit{Relating } h_X \textit{ to } N_X \textit{ and the convergence of the smoother}}$

   Theorem 191 derives the smoother error estimate 5.50:

   $$|s_e(x)-f_d(x)|\leq|f_d|_{w,\theta}\left(1+K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\,(c_{\Omega,\theta}h_X)^{\eta+\delta_G}+\sqrt{N_X\rho}\right),$$

   when $h_X\leq\min\{h_{\Omega,\theta},r_G\}$; with the awkward term N_X which is linked to h_X in an indirect manner. In order to have a meaningful concept of the convergence of a smoother a relationship between h_X and N_X needs to be introduced. This situation did not arise with interpolant.

2. *In general the data X is scattered and for a given value of h_X the number of points N_X can be arbitrarily large. However, for data on a regular, rectangular grid we have the relation*

   $$d^{d/2}\,\mathrm{vol}(grid)=N_X(h_X)^d. \tag{5.56}$$

3. *In Williams [22] several* ***1-dimensional*** *numerical experiments were run to compare the convergence of the* ***zero order*** *Exact smoother with the predicted convergence. One-dimensional data sets were constructed using a uniform distribution on the interval $\Omega=[-1.5,1.5]$. Each of 20 data files were exponentially sampled using a multiplier of approximately 1.2 and a maximum of 5000 points, and then $\log_{10}h_X$ was plotted against $\log_{10}N$ where $N=|X|$. It then seemed quite reasonable to use a least-squares linear fit and in this case we obtained the relation*

   $$h_X\simeq 3.09N^{-0.81}. \tag{5.57}$$

   For ease of calculation let

   $$h_X=h_1(N_X)^{-a}, \quad h_1=3.09, \quad a=0.81. \tag{5.58}$$

4. *A barrier to the use of such a formula as 5.57 in higher dimensions is the difficulty of actually calculating h_X for a given data set. If a sequence of independent test data sets was generated by a uniform distribution in each dimension then the constants a and h_1 might be defined as the upper bound of the confidence interval of a statistical distribution. Also, noting regular grid formula 5.56 we might hypothesize a relationship of the form*

$$h_X = h_d \left(N_X\right)^{-a_d d},$$

*for **higher dimensions**.*

5. $\boxed{\text{The order of convergence}}$ *Assuming 5.58 we will now show that for a sequence of independent data points X_k of increasing density there exists a sequence ρ_k of smoothing coefficients such that the smoother error is of order $\eta_G = \eta + \delta_G$ in h_{X_k}.*

By Theorem 191

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_{X_k}\right)^{\eta_G} + \sqrt{N_{X_k} \rho_k}\right), \quad x \in \overline{\Omega},$$

when $h_{X_k} < \min\left\{h_{\Omega,\theta}, r_G\right\}$. For clarity define the constants

$$A = |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta}\right)^{\eta_G}, \quad B = |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right), \tag{5.59}$$

so that

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \leq A \left(h_{X_k}\right)^{\eta_G} + B\sqrt{\rho_k}\sqrt{N_{X_k}}, \quad x \in \overline{\Omega}.$$

Then condition 5.58 implies

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \leq A \left(h_{X_k}\right)^{\eta_G} + B\sqrt{\rho_k} \left(\frac{h_{X_k}}{h_1}\right)^{-\frac{1}{2a}}, \quad x \in \overline{\Omega}, \tag{5.60}$$

and we want to minimize the right side as a function of h_{X_k}. This is easily done by setting the derivative to zero so that

$$\begin{aligned} D_x \left(A x^{\eta_G} + B\sqrt{\rho_k} \left(\frac{x}{h_1}\right)^{-\frac{1}{2a}} \right)\Bigg|_{x = h_{X_k}} &= A\eta_G \left(h_{X_k}\right)^{\eta_G - 1} - \frac{1}{2a h_1} B\sqrt{\rho_k} \left(\frac{h_{X_k}}{h_1}\right)^{-\frac{1}{2a} - 1} \\ &= 0, \end{aligned}$$

and a unique minimum is obtained when

$$\sqrt{\rho_k} = \frac{A}{B} \frac{2a\eta_G}{\left(h_1\right)^{\frac{1}{2a}}} \left(h_{X_k}\right)^{\eta_G + \frac{1}{2a}}. \tag{5.61}$$

*By substituting for $\sqrt{\rho_k}$ in the right side of 5.60 we find that the corresponding **minimum error** value is given by*

$$\begin{aligned} \left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| &\leq A \left(h_{X_k}\right)^{\eta_G} + B\sqrt{\rho_k} \left(\frac{h_{X_k}}{h_1}\right)^{-\frac{1}{2a}} \\ &= A \left(h_{X_k}\right)^{\eta_G} + B \left(\frac{A}{B} \frac{2a\eta_G}{\left(h_1\right)^{\frac{1}{2a}}} \left(h_{X_k}\right)^{\eta_G + \frac{1}{2a}}\right) \left(\frac{h_{X_k}}{h_1}\right)^{-\frac{1}{2a}} \\ &= A \left(h_{X_k}\right)^{\eta_G} + 2a\eta_G A \left(h_{X_k}\right)^{\eta_G} \\ &= \left(1 + 2a\eta_G\right) A \left(h_{X_k}\right)^{\eta_G} \\ &= \left(1 + 2a\eta_G\right) |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta} h_{X_k}\right)^{\eta_G}. \end{aligned}$$

Thus if $h_{X_k} \to 0$ and ρ_k is chosen to satisfy 5.61 then the Exact smoother error always converges uniformly to zero and the order of convergence is at least $\left(h_{X_k}\right)^{\eta + \delta_G}$. Note also that the minimum error is independent of h_1.

6. *From 5.61 the convergence of ρ_k to zero is of order $2\eta_G + \frac{1}{a}$ in h_X. One might consider saying that the smaller this order is the better in order to avoid stability problems when ρ is small?*

7. $\boxed{\text{Error behavior in terms of } \rho}$ *The ρ error curve is well known, always becoming constant with large ρ and decreasing as ρ decreases and then it may increase again to the interpolant error value. Its observed slope behavior depends on the basis function: with decreasing ρ:*

   - ***Thin-plate spline*** *slope increases like 5.50 then levels out.*
   - ***Shifted thin-plate spline*** *slope keeps increasing like 1.26.*
   - ***Gaussian*** *slope oscillates with increasing positive amplitude.*
   - ***B-splines*** *- when $n = l = 1$ the slope tends to zero. For other splines it doe not.*

   Note that observing the behavior of the smoother for small ρ may run into stability problems and the limitations of accuracy.

   If we assume our X data lies in $\mathbb{R}^1$ and that h_X and N_X satisfy 5.58 then we obtain the error estimate 5.60:

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \leq A\left(h_{X_k}\right)^{\eta_G} + B\sqrt{\rho}\left(\frac{h_{X_k}}{h_1}\right)^{-\frac{1}{2a}}, \quad x \in \overline{\Omega}.$$

   Clearly the slope always tends to infinity for small ρ and always tends to infinity for large ρ, which leaves much to be desired. However, in Subsection 5.8.3 we will establish error estimates that, for given X, are constant for large ρ.

5.8.2 Doubled order of convergence

In this subsection a double rate of convergence in terms of h_X will be demonstrated for data functions that are linear combinations of Riesz representers. We will again use the assumptions and notation of Lemma 188 regarding the open set which contains the independent data points. The convergence results will be applied to the case of semi-regular sequences of independent data points.

The operator $\mathcal{L}_X$ of Definition 168 is not onto but by Theorem 170 it is 1-1, and since $\mathcal{L}_X\left(X^\theta_w\right)$ is not dense in V the usual definition of the adjoint does not apply to $\mathcal{L}_X^{-1}$. However, by Theorem 175, $\mathcal{L}_X^*\mathcal{L}_X : X^\theta_w \to X^\theta_w$ is 1-1 and onto and this property is used in the next lemma to extend $\mathcal{L}_X^{-1}$ to V as a continuous operator and allows a definition of $\left(\mathcal{L}_X^{-1}\right)^*$:

Lemma 195 ***The operator*** $\left(\mathcal{L}_X^{-1}\right)^*$*: The operator $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^* : V \to X^\theta_w$ is a continuous extension of $\mathcal{L}_X^{-1}$ to V. Define the operator $\left(\mathcal{L}_X^{-1}\right)^*$ by $\left(\mathcal{L}_X^{-1}\right)^* = \left(\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*\right)^*$. Then $\left(\mathcal{L}_X^{-1}\right)^*$ has the following properties:*

1. $\left(\mathcal{L}_X^{-1}\right)^* = \mathcal{L}_X\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}$.
2. $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1} = \mathcal{L}_X^{-1}\left(\mathcal{L}_X^{-1}\right)^*$.
3. $\mathcal{L}_X^*\left(\mathcal{L}_X^{-1}\right)^* = I$.
4. $\left(\mathcal{L}_X^{-1}\right)^{**} = \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*$.

Proof. The operator $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^* : X^\theta_w \to V$ is a continuous extension of $\mathcal{L}_X^{-1}$ to V because $\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*\mathcal{L}_X = I$.

Part 1 Since $\mathcal{L}_X^*\mathcal{L}_X : X^\theta_w \to X^\theta_w$ is 1-1 and onto, $\left(\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\right)^* = \left(\left(\mathcal{L}_X^*\mathcal{L}_X\right)^*\right)^{-1} = \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}$ and so, $\left(\mathcal{L}_X^{-1}\right)^* = \left(\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*\right)^* = \mathcal{L}_X\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}$.

Part 2 Left-compose $\mathcal{L}_X^{-1}$ with the equation of part 1.

Part 3 Left-compose $\mathcal{L}_X^*$ with the equation of part 1.

Part 4 By definition $\left(\mathcal{L}_X^{-1}\right)^{**} = \left(\left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*\right)^{**} = \left(\mathcal{L}_X^*\mathcal{L}_X\right)^{-1}\mathcal{L}_X^*$. ■

The next result gives more properties of the function v_x studied in the previous subsection, as well as defining and studying the object $r_{V,x'} = \left(\mathcal{L}_X^{-1}\right)^* r_{x'} \in V$ which is used to derive the double order convergence estimates.

Theorem 196 *Suppose $\mathcal{S}_X^e$ is the Exact smoother mapping and $f_d \in X_w^\theta$ is a data function. Then:*

1. $\mathcal{L}_X^* R_{V,x} = R_x$ *and* $R_{V,x} = \left(\mathcal{L}_X^{-1}\right)^* R_x$.
2. $f_d(x) - \mathcal{S}_X^e f_d(x) = \rho \left(\left(\mathcal{L}_X^{-1}\right)^* \mathcal{Q} f_d, R_{V,x} \right)_V = \langle f_d, v_x \rangle_{w,\theta}$.
3. $R_{x'}(x) - \mathcal{S}_X^e R_{x'}(x) = \rho \left(r_{V,x'}, R_{V,x} \right)_V = \mathcal{Q}_{x'} \overline{v_{x'}(x)}$, *where* $r_{V,x'} = \left(\mathcal{L}_X^{-1}\right)^* r_{x'}$ *and the subscript on the operator* $\mathcal{Q}_{x'}$ *indicates the action variable.*
4. $\rho \left\| r_{V,x} \right\|_V^2 \leq r_x(x)$.

Proof. Part 1 From 5.28 and 5.29, $R_x = \mathcal{L}_X^* R_{V,x}$, $\mathfrak{R}_{V,x} = \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} R_x$ and $R_{V,x} = \mathcal{L}_X \mathfrak{R}_{V,x}$. Thus by part 1 Lemma 195, $R_{V,x} = \mathcal{L}_X \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} R_x = \left(\mathcal{L}_X^{-1}\right)^* R_x$.

Part 2 From 5.11, $\mathcal{S}_X^e f_d = f_d - \rho \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \mathcal{Q} f_d$. Hence

$$\begin{aligned} f_d(x) - \mathcal{S}_X^e f_d(x) = \left(\mathcal{L}_X \left(I - \mathcal{S}_X^e\right) f_d, R_{V,x} \right)_V &= \rho \left(\mathcal{L}_X \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \mathcal{Q} f_d, R_{V,x} \right)_V \\ &= \rho \left(\left(\mathcal{L}_X^{-1}\right)^* \mathcal{Q} f_d, \mathcal{L}_X^* R_{V,x} \right)_{w,\theta} \\ &= \rho \left\langle f_d, \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \mathcal{L}_X^* R_{V,x} \right\rangle_{w,\theta}. \end{aligned}$$

But the equations 5.29 imply $\mathfrak{R}_{V,x} = \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \mathcal{L}_X^* R_{V,x}$ and by 5.35, $\rho \mathfrak{R}_{V,x} = v_x$ so it follows that

$$f_d(x) - \mathcal{S}_X^e f_d(x) = \rho \left\langle f_d, \mathfrak{R}_{V,x} \right\rangle_{w,\theta} = \langle f_d, v_x \rangle_{w,\theta}.$$

Part 3 Substitute $f_d = R_{x'}$ in part 2 so that

$$R_{x'}(x) - \mathcal{S}_X^e R_{x'}(x) = \rho \left(r_{V,x'}, R_{V,x} \right)_V = \mathcal{Q} v_x(x') = \mathcal{Q}_{x'} \overline{v_{x'}(x)}.$$

Part 4 From part 3, $r_{V,x} = \left(\mathcal{L}_X^{-1}\right)^* r_x$ so that from part 1 Lemma 195

$$\begin{aligned} \left\| r_{V,x} \right\|_V^2 = \left(r_{V,x}, r_{V,x} \right)_V = \left(\left(\mathcal{L}_X^{-1}\right)^* r_x, \left(\mathcal{L}_X^{-1}\right)^* r_x \right)_V &= \left(\mathcal{L}_X \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} r_x, \mathcal{L}_X \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} r_x \right)_V \\ &= \left(\left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} r_x, r_x \right)_{w,\theta}. \end{aligned}$$

Now let $z_x = \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} r_x$ so that Theorem 116 and part 7 Theorem 125 imply

$$\left\| r_{V,x} \right\|_V^2 = \mathcal{Q} z_x(x). \tag{5.62}$$

But by part 4 Theorem 170

$$\mathcal{L}_X^* \mathcal{L}_X z_x = \rho \mathcal{Q} z_x + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X z_x = r_x,$$

which is equivalent to the system

$$\mathcal{Q} w_x + \widetilde{\mathcal{E}}_X^* \beta_x = r_x, \tag{5.63}$$

$$\frac{1}{N\rho} \widetilde{\mathcal{E}}_X w_x = \beta_x, \tag{5.64}$$

where

$$\rho z_x = w_x. \tag{5.65}$$

We now want to calculate $\left\| r_{V,x} \right\|_V$ by solving the system 5.63, 5.64, 5.65 and then using 5.62. To do this we use the analogue of the method used to calculate $\left\| R_{V,x} \right\|_V = \mathfrak{R}_{V,x}(x)$ subsequent to Theorem 186. This involved solving the system 5.33, 5.34, 5.35 for $\mathfrak{R}_{V,x}$. Using this technique yields

$$w_x(y) = r_x(y) - r_{X_2,x}^T \left(N\rho \left(I + L_{X_2} L_{X_2}^T \right) + \overline{r_{X_2,X_2}} \right)^{-1} \overline{r_{X_2,y}},$$

so that

$$\begin{aligned}\rho\left\|r_{V,x}\right\|_V^2 &= \rho\mathcal{Q}z_x(x)=\mathcal{Q}w_x(x)=w_x(x)\\ &= r_x(x)-r_{X_2,x}^T\left(N\rho\ \left(I+L_{X_2}L_{X_2}^T\right)+\overline{r_{X_2,X_2}}\right)^{-1}\overline{r_{X_2,x}}\\ &\le r_x(x),\end{aligned}$$

since it was shown after 5.38 that $\left(N\rho\ \left(I+L_{X_2}L_{X_2}^T\right)+\overline{r_{X_2,X_2}}\right)^{-1}$ is positive definite. ■
We can now derive our double order convergence estimates.

Theorem 197 *Suppose $\mathcal{S}_X^e$ is the Exact smoother mapping of Definition 146, $s_e=\mathcal{S}_X^e f_d$ and R_x is the Riesz representer of the functional $f\to f(x)$ defined using a minimal unisolvent subset A. Set $\eta_G=\eta+\delta_G$. Then:*

1. *If $x,x'\in\overline{\Omega}$ and $h_X<\min\{h_{\Omega,\theta},r_G\}$ we have the double order of convergence estimate*

$$\left|R_{x'}(x)-\mathcal{S}_X^e R_{x'}(x)\right|\le\left(1+K'_{\Omega,\theta}\right)^2\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}+\sqrt{N\rho}\right).$$

2. *If $X'=\{x'_k\}_{k=1}^{N'}\subset\Omega$, $h_X<\min\{h_{\Omega,\theta},r_G\}$ and the data function has the form $f_d=\sum\limits_{k=1}^{N'}\beta_k R_{x'_k}$ for some $\beta_k\in\mathbb{C}$ then*

$$\left|f_d(x)-s_e(x)\right|\le\left(\sum_{k=1}^{N'}\left|\beta_k\right|\right)\left(1+K'_{\Omega,\theta}\right)^2\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}+\sqrt{N\rho}\right),\quad x\in\overline{\Omega}.$$

3. *If $A\subset X'$ then the data function of part 2 lies in $W_{G,X'}$.*

Proof. Part 1 If $x,x'\in\mathbb{R}^d$ then by first using the Cauchy-Schwartz inequality and then the estimates 5.41 and part 4 Theorem 196 for $\left\|R_{V,x}\right\|_V$ and $\left\|r_{V,x'}\right\|_V$ we obtain

$$\begin{aligned}\left|R_{x'}(x)-\mathcal{S}_X^e R_{x'}(x)\right| &= \rho\left|\left(r_{V,x'},R_{V,x}\right)_V\right|\le\rho\left\|r_{V,x'}\right\|_V\left\|R_{V,x}\right\|_V\\ &\le \rho\sqrt{\frac{r_{x'}(x')}{\rho}}\sqrt{\frac{r_x(x)+N\rho\left|\widetilde{l}(x)\right|^2}{\rho}}\\ &\le \sqrt{r_{x'}(x')}\sqrt{r_x(x)+N\rho\left|\widetilde{l}(x)\right|^2}\\ &\le \sqrt{r_{x'}(x')}\left(\sqrt{r_x(x)}+\sqrt{N\rho}\left|\widetilde{l}(x)\right|\right)\\ &\le \sqrt{r_{x'}(x')}\left(\sqrt{r_x(x)}+\sqrt{N\rho}\left|\widetilde{l}(x)\right|_1\right).\end{aligned}\tag{5.66}$$

Regarding the last inequality, if $x,x'\in\Omega$ and $h_X<h_{\Omega,\theta}$ then we can uniformly estimate $\sqrt{r_{x'}(x')}$ and $\sqrt{r_x(x)}$using 5.53 and uniformly estimate $\left|\widetilde{l}(x')\right|_1$ using 5.42. Hence

$$\begin{aligned}\left|R_{x'}(x)-\mathcal{S}_X^e R_{x'}(x)\right| &\le \sqrt{r_{x'}(x')}\left(\sqrt{r_x(x)}+K'_{\Omega,\theta}\sqrt{N\rho}\right)\\ &\le \left(1+K'_{\Omega,\theta}\right)\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}\left(\left(1+K'_{\Omega,\theta}\right)\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}+K'_{\Omega,\theta}\sqrt{N\rho}\right)\\ &= \left(1+K'_{\Omega,\theta}\right)^2\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}+\sqrt{N\rho}\right).\end{aligned}$$

Part 2 The proof proceeds by applying the uniform estimate of part 1 to each term of $f_d = \sum\limits_{k=1}^{N'} \beta_k R_{x'_k}$. Thus

$$
\begin{aligned}
\left|f_d(x) - \mathcal{S}_X^e f_d(x)\right| &\leq \left|\sum_{k=1}^{N'} \beta_k \left(R_{x'_k}(x) - \mathcal{S}_X^e R_{x'_k}(x)\right)\right| \\
&\leq \left(\sum_{k=1}^{N'} |\beta_k|\right) \max_{\substack{x\in\overline{\Omega} \\ 1\leq k\leq N'}} \left|R_{x'_k}(x) - \mathcal{S}_X^e R_{x'_k}(x)\right| \\
&\leq \left(\sum_{k=1}^{N'} |\beta_k|\right) \left(1 + K'_{\Omega,\theta}\right)^2 \sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} + \sqrt{N\rho}\right).
\end{aligned}
$$

Part 3 If $A \subset X'$ then by part 6 Theorem 166, $R_{x'_k} \in W_{G,X'}$ for each k, and so $f_d \in W_{G,X'}$. ■

Remark 198 *Following the approach of Remark 194 we will now show that if we assume h_X and N are related by 5.58 then for a sequence of independent data sets X_k in $\mathbb{R}^1$ there exists a sequence ρ_k of smoothing coefficients such that the smoother error is of order $2\eta_G$ in h_{X_k}.*

By part 1 Theorem 197, when $h_{X_k} < \min\{h_{\Omega,\theta}, r_G\}$

$$
\left|f_d(x) - \mathcal{S}_{X_k}^e f_d(x)\right| \leq \left(1 + K'_{\Omega,\theta}\right)^2 \sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} \left(\sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} + \sqrt{N\rho}\right), \quad x \in \overline{\Omega}.
$$

For clarity define the constants

$$
A = \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta}\right)^{\eta_G}, \quad B = 1 + K'_{\Omega,\theta}, \tag{5.67}
$$

so that

$$
\left|f_d(x) - \mathcal{S}_{X_k}^e f_d(x)\right| \leq A^2 \left(h_{X_k}\right)^{2\eta_G} + AB\sqrt{\rho} \left(h_{X_k}\right)^{\eta_G} \sqrt{N_{X_k}}, \quad x \in \overline{\Omega}.
$$

The condition 5.58 implies

$$
\left|f_d(x) - \mathcal{S}_{X_k}^e f_d(x)\right| \leq A^2 \left(h_{X_k}\right)^{2\eta_G} + AB\sqrt{\rho} \left(h_1\right)^{\frac{1}{2a}} \left(h_{X_k}\right)^{\eta_G - \frac{1}{2a}}, \quad x \in \overline{\Omega}, \tag{5.68}
$$

and we want to minimize the right side as a function of h_{X_k}. This is easily done by setting the derivative to zero so that

$$
\begin{aligned}
D_{h_{X_k}} \left(A^2 \left(h_{X_k}\right)^{2\eta_G} + AB\sqrt{\rho} \left(h_1\right)^{\frac{1}{2a}} \left(h_{X_k}\right)^{\eta_G - \frac{1}{2a}}\right) &= 2\eta_G A^2 \left(h_{X_k}\right)^{2\eta_G - 1} + \\
&\quad + \left(\eta_G - \frac{1}{2a}\right) AB\sqrt{\rho} \left(h_{X_k}\right)^{\eta_G - \frac{1}{2a} - 1} \\
&= 0,
\end{aligned}
$$

and a unique minimum is obtained if $\eta_G < \frac{1}{2a}$. In this case h_{X_k} satisfies

$$
2\eta_G A \left(h_{X_k}\right)^{\eta_G + \frac{1}{2a}} = \left(\frac{1}{2a} - \eta_G\right) B\sqrt{\rho}, \tag{5.69}
$$

which can be written

$$
\sqrt{\rho_k} = \frac{2\eta_G}{\frac{1}{2a} - \eta_G} \frac{A}{B} \left(h_{X_k}\right)^{\eta_G + \frac{1}{2a}}. \tag{5.70}
$$

By substituting for ρ in the right side of 5.68 we find that the corresponding minimum error value is given by

$$
\begin{aligned}
A^2 \left(h_{X_k}\right)^{2\eta_G} + AB\sqrt{\rho} \left(h_1\right)^{\frac{1}{2a}} \left(h_{X_k}\right)^{\eta_G - \frac{1}{2a}} &= A^2 \left(h_{X_k}\right)^{2\eta_G} + \\
&\quad + AB \left(\frac{2\eta_G}{\frac{1}{2a} - \eta_G} \frac{A}{B} \left(h_{X_k}\right)^{\eta_G + \frac{1}{2a}}\right) \left(h_{X_k}\right)^{\eta_G - \frac{1}{2a}} \\
&= A^2 \left(h_{X_k}\right)^{2\eta_G} + A^2 \frac{2\eta_G}{\frac{1}{2a} - \eta_G} \left(h_{X_k}\right)^{2\eta_G} \\
&= \frac{\frac{1}{2a} + \eta_G}{\frac{1}{2a} - \eta_G} A^2 \left(h_{X_k}\right)^{2\eta_G},
\end{aligned}
$$

so that

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \le \frac{\frac{1}{2a} + \eta_G}{\frac{1}{2a} - \eta_G} A^2 (h_{X_k})^{2\eta_G}, \quad x \in \overline{\Omega}.$$

Thus if $\eta_G < \frac{1}{2a}$ *and* $h_{X_k} \to 0$ *and* ρ_k *is chosen to satisfy 5.70, then the Exact smoother error will converge uniformly to zero and the order of convergence will be at least* $(h_{X_k})^{2\eta_G}$.

On the other hand if $\eta_G \ge \frac{1}{2a}$ *then noting 5.68 choose* ρ_k *so that* $\sqrt{\rho_k}\,(h_{X_k})^{\eta_G - \frac{1}{2a}} = (h_{X_k})^{2\eta_G}$ *i.e. so that*

$$\sqrt{\rho_k} = (h_{X_k})^{\eta_G + \frac{1}{2a}},$$

and hence 5.68 becomes

$$\left|f_d(x) - \mathcal{S}^e_{X_k} f_d(x)\right| \le \left(A^2 + AB\,(h_1)^{\frac{1}{2a}}\right)(h_{X_k})^{2\eta_G}, \quad x \in \overline{\Omega},$$

and again the Exact smoother error converges uniformly to zero and the order of convergence is at least $(h_{X_k})^{2\eta_G}$.

5.8.3 Error estimates bounded in the smoothing parameter

As a function of the smoothing parameter ρ the convergence estimates of the previous two subsections tend to infinity as $\rho \to \infty$. However, numerical experiments indicate that the error is a bounded function of ρ and we will now derive some convergence estimates with this property. The first step is to derive an estimate for the Exact smoother error which involves the Exact smoother functional J_e. First recall the notation of Lemma 188.

Theorem 199 *Let* w *be a weight function with properties W2 and W3 for order* θ *and parameter* κ *and set* $\eta = \max\limits_{\kappa} \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. *Assume* G *is a basis function of order* θ *such that there exist constants* $c_G > 0$ *and* $\delta_G \ge 0$ *such that 5.49 holds. Set* $\eta_G = \eta + \delta_G$.

Suppose s_e *is the Exact smoother of the data function* f_d *on the unisolvent independent data* X *with* N_X *points. Suppose* $A \subset X$ *is a minimal* θ*-unisolvent set and that* $\{l_k\}_{k=1}^M$ *is the corresponding unique cardinal basis for* P_θ. *Set* $A_x = A \cup \{x\}$.

Then if $x \in \overline{\Omega}$ *and* $\operatorname{diam} A_x \le r_G$ *then*

$$|f_d(x) - s_e(x)| \le |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\,(\operatorname{diam} A_x)^{\eta_G} + K'_{\Omega,\theta}\sqrt{N_X J_e[s_e]}, \tag{5.71}$$

and further if $h_X \le \{h_{\Omega,\theta}, r_G\}$ *then*

$$|f_d(x) - s_e(x)| \le |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\,(c_{\Omega,\theta} h_X)^{\eta_G} + K'_{\Omega,\theta}\sqrt{N_X J_e[s_e]}. \tag{5.72}$$

Here J_e *is the Exact smoother functional.*

Proof. Fix $x \in \Omega$ and suppose $A = \{a_k\}_{k=1}^M$. Then

$$f_d(x) - s_e(x) = (f_d - s_e, R_x)_{w,\theta} = \langle f_d - s_e, r_x\rangle_{w,\theta} + \sum_{k=1}^M (f_d(a_k) - s_e(a_k))\, l_k(x),$$

so that

$$\begin{aligned}
|f_d(x) - s_e(x)| &\le \left|\langle f_d - s_e, r_x\rangle_{w,\theta}\right| + \sum_{k=1}^M |f_d(a_k) - s_e(a_k)|\,|l_k(x)| \\
&\le |f_d - s_e|_{w,\theta}\,|r_x|_{w,\theta} + \left(\max_k |f_d(a_k) - s_e(a_k)|\right)\sum_{k=1}^M |l_k(x)| \\
&= |f_d - s_e|_{w,\theta}\sqrt{r_x(x)} + \left(\max_k |f_d(a_k) - s_e(a_k)|\right)\sum_{k=1}^M |l_k(x)|.
\end{aligned}$$

Since $x \in \Omega$ the estimates 5.42 and 5.52 imply

$$|f_d(x) - s_e(x)| \le |f_d - s_a|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\,(\operatorname{diam} A_x)^{\eta_G} + K'_{\Omega,\theta}\left(\max_k |f_d(a_k) - s_e(a_k)|\right).$$

From 5.29, $f_d(x) - s_e(x) = (\mathcal{L}_X(f_d - s_e), R_{V,x})_V$ so that

$$|f_d(a_k) - s_e(a_k)| \leq \|\mathcal{L}_X(f_d - s_e)\|_V \|R_{V,a_k}\|_V,$$

but Remark 169 with $f = s_e$ and part 3 of Theorem 171 with $f = f_d$ yields $\|\mathcal{L}_X(f_d - s_e)\|_V \leq \sqrt{J_e[s_e]}$ so that

$$|f_d(x) - s_e(x)| \leq |f_d - s_e|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta_G} + K'_{\Omega,\theta} \sqrt{J_e[s_e]} \max_k \|R_{V,a_k}\|_V.$$

From 5.41, $\|R_{V,a_k}\|_V \leq \sqrt{N_X}$ and by 5.8 with $f = f_d$, $|f_d - s_e|_{w,\theta} \leq |f_d|_{w,\theta}$ so that

$$|f_d(x) - s_e(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta_G} + K'_{\Omega,\theta} \sqrt{N_X J_e[s_e]}, \tag{5.73}$$

which proves 5.71. Finally, since $h_X \leq \{h_{\Omega,\theta}, r_G\}$, Lemma 188 implies $\operatorname{diam} A_x \leq c_{\Omega,\theta} h_X$ and so 5.73 implies 5.72. ∎

Next we estimate the term $J_e[s_e]$ which occurs in both inequalities of the last theorem.

Theorem 200 *Let w be a weight function with properties W2 and W3 for order θ and parameter κ and set $\eta = \max_\kappa \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Assume G is a basis function of order θ such that there exist constants $c_G > 0$ and $\delta_G \geq 0$ such that 5.49 holds. Set $\eta_G = \eta + \delta_G$.*

Suppose s_e is the Exact smoother of the data function f_d on the independent data X.

Then if Ω is a bounded, open, connected subset of $\mathbb{R}^d$ having the cone property:

$$\sqrt{J_e[s_e]} \leq |f_d|_{w,\theta} \min\left\{\sqrt{\rho}, \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} \Omega\right)^{\eta_G}\right\}, \tag{5.74}$$

where $\rho > 0$ is the smoothing parameter.

Proof. Choose a unisolvent set $A \subset X$ to define the operators $\mathcal{P}$ and $\mathcal{Q}$.

From the definition of the smoother problem $J_e[s_e] \leq J_e[f_d] = \rho |f_d|^2_{w,\theta}$. Also

$$J_e[s_e] \leq J_e[\mathcal{P} f_d] = \rho |\mathcal{P} f_d|^2_{w,\theta} + \frac{1}{N} \sum_{k=1}^{N} \left|\mathcal{P} f_d\left(x^{(k)}\right) - f_d\left(x^{(k)}\right)\right|^2 = \frac{1}{N} \sum_{k=1}^{N} \left|\mathcal{Q} f_d\left(x^{(k)}\right)\right|^2 \leq \max_{x \in \overline{\Omega}} |\mathcal{Q} f_d(x)|^2.$$

Fix $x \in \overline{\Omega}$ and using the properties of $\mathcal{Q}$ given in Theorem 116 and the properties of the semi-Riesz representer r_x given in Theorem 125 we have

$$|\mathcal{Q} f_d(x)| = \left|\langle f_d, r_x \rangle_{w,\theta}\right| \leq |f_d|_{w,\theta} |r_x|_{w,\theta} = |f_d|_{w,\theta} \sqrt{r_x(x)}.$$

But from 5.49

$$\sqrt{r_x(x)} \leq \left(1 + \left|\widetilde{l}(x)\right|_1\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta_G} \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} \Omega\right)^{\eta_G},$$

so that

$$\sqrt{J_e[s_e]} \leq \max_{x \in \overline{\Omega}} |\mathcal{Q} f_d(x)| \leq |f_d|_{w,\theta} \max_{x \in \overline{\Omega}} \sqrt{r_x(x)} \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} \Omega\right)^{\eta_G}.$$

∎

We now combine the last two results to obtain our improved error estimates which are bounded in ρ.

Theorem 201 *Let w be a weight function with properties W2 and W3 for order θ and parameter κ and set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Assume G is a basis function of order θ such that there exist constants $c_G > 0$ and $\delta_G \geq 0$ such that 5.49 holds. Set $\eta_G = \eta + \delta_G$.*

Suppose s_e is the Exact smoother of the data function f_d on the unisolvent independent data X with N_X points. Suppose the notation and assumptions of Lemma 188 hold so that the data point density is $h_X = \sup_{\omega \in \Omega} \operatorname{dist}(\omega; X)$.

Then there exist constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta} > 0$ such that

$$\begin{aligned} |f_d(x) - s_e(x)| \leq{} & |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} A_x\right)^{\eta_G} + \\ & + |f_d|_{w,\theta} \sqrt{N_X} \min\left\{\sqrt{\rho}, \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam} \Omega\right)^{\eta_G}\right\}, \end{aligned} \tag{5.75}$$

and

$$\begin{aligned} |f_d(x) - s_e(x)| \quad &\leq \quad |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} + \\ &+ |f_d|_{w,\theta} K'_{\Omega,\theta} \sqrt{N_X} \min\left\{\sqrt{\rho}, \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_G} \left(\operatorname{diam}\Omega\right)^{\eta_G}\right\}, \end{aligned}$$

for $x \in \Omega$.

Here the constants $c_{\Omega,\theta}, h_{\Omega,\theta}, K'_{\Omega,\theta}$ *only depend on* Ω, θ, κ *and* d.

Proof. The inequalities of this theorem are obtained by applying the estimates for $\sqrt{J_e[s_e]}$ proved in Theorem 200 to the estimates of Theorem 199. ■

Remark 202 *For a given independent data set* X *the bound 5.75 is a bounded function of* ρ. *However, numerical experiments indicate that there should be a bound that is also independent of* N_X.

6

The Approximate smoother and its convergence to the Exact smoother and the data function

6.1 Introduction

The Approximate smoother problem is derived from the Exact smoother problem by restricting the range of the minimizing functions from X_w^θ to a space $W_{G,X'}$, where X' is an arbitrary set of unisolvent points in $\mathbb{R}^d$.

The Approximate smoother problem is solved twice, first using Hilbert space methods and then using matrix methods. A matrix equation is derived and the construction and solution of this equation is shown to be scalable w.r.t. the number of data points i.e. it depends linearly on the number of data points.

Estimates are first derived for the pointwise convergence of the Approximate smoother to the Exact smoother and these are then added to the Exact smoother error estimates to obtain the error of the Approximate smoother. These estimates involve uniform pointwise convergence and derive orders of convergence. However they are basically unsatisfactory and are very rough estimates when compared with numerical results. I have not included the results of any numerical experiments in this chapter.

Section by section:

Section 6.2 A study of the convolution spaces $J_G = G * \widehat{S}_{\emptyset,\theta} \oplus P_\theta$ and $\dot{J}_G = G * \widehat{S}_{\emptyset,\theta}$. The hat indicates the Fourier transform.

Section 6.3 Formulation of the Approximate smoothing problem - The Approximate smoother problem is derived from the Exact smoother problem studied in Chapter 5 by restricting the range of the minimizing functions from X_w^θ to a space $W_{G,X'}$, where X' is an arbitrary set of θ-unisolvent points in $\mathbb{R}^d$. The first step is to assume that X' is a regular grid containing the Exact smoother data and to approximate or discretize the functions in X_w^θ using this grid. This leads to the finite dimensional subspace $W_{G,X'}$ and the Approximate smoother problem, namely $\min\limits_{f \in W_{G,X'}} J_e[f]$, where J_e is the Exact smoothing functional. We then generalize this problem to an arbitrary unisolvent X'.

Section 6.4 Studying the Approximate smoothing problem using Hilbert space techniques. We now know the Approximate smoother exists, is unique and is a member of $W_{G,X'}$. The next step is to derive a matrix equation for the coefficients of the X'-translated basis functions and the basis polynomials. This proof is quite similar to that of Theorem 183 which derives the Exact smoother matrix equation.

Hilbert space techniques are used to show that the Approximate smoothing problem has a unique solution. We obtain a matrix equation for the Approximate smoother.

Section 6.5 Matrix techniques are used to derive the matrix equation. We write $f = \sum\limits_{i=1}^{N'} \alpha_i G(\cdot - x_i') + \sum\limits_{j=1}^{M} \beta_j p_j \in W_{G,X'}$ and calculate $J[f]$ as a quadratic form in terms of $\begin{pmatrix} \alpha^T & \beta^T \end{pmatrix}$ restrained by $P_{X'}^T \alpha = 0$. We next show the quadratic component is is positive definite and then obtain the matrix equation for the Exact smoother using Lagrange multipliers.

Section 6.6 For a bounded data region Ω we first derive some uniform, pointwise convergence results which do not involve orders of convergence e.g. in Corollary 229 it is shown that as the points in X' get closer to those in X the Approximate smoother converges uniformly to the Exact smoother on $\overline{\Omega}$. We next derive some orders of pointwise convergence for the Approximate smoother to the Exact smoother and then add these to the Exact smoother estimates of Chapter 5 to obtain the Approximate smoother error.

6.2 The convolution spaces $\dot{J}_G$ and J_G

In this section we define the convolution spaces J_G which will be used to discretize the space X_w^θ in the derivation of the Approximate smoothing problem in Section 6.3. The notation J_G comes from Section 3 of Nira Dyn's review paper [4] and similar results are proved here.

Theorem 203 *Suppose the weight function w has property W2, and suppose G is a basis distribution of order $\theta \geq 1$ generated by w. Then $\varphi \in \widehat{S}_{\emptyset,\theta}$ implies $G * \varphi \in X_w^\theta$ and if $\psi \in \widehat{S}_{\emptyset,\theta}$*

$$\langle G * \varphi, G * \psi \rangle_{w,\theta} = \int \frac{\widehat{\varphi}\overline{\widehat{\psi}}}{w\left|\cdot\right|^{2\theta}} = \left[G, \varphi_{\_} * \overline{\psi}\right], \tag{6.1}$$

where $\varphi_{\_}(x) = \varphi(-x)$. Also, if $f \in X_w^\theta$ then

$$\langle f, G * \varphi \rangle_{w,\theta} = \left[f, \overline{\varphi}\right]. \tag{6.2}$$

Proof. Corollary 114 implies that $\varphi \in \widehat{S}_{\emptyset,\theta}$ implies $G * \varphi \in X_w^\theta$ for $\theta \geq 1$ and that 6.2 is true for $\theta \geq 1$. It remains to prove 6.1.

From Corollary 114, $|G * \varphi|_{w,\theta}^2 = \int \frac{|\widehat{\varphi}|^2}{w|\cdot|^{2\theta}}$ when $\varphi \in \widehat{S}_{\emptyset,\theta}$. Since X_w^θ is a semi-inner product vector space with complex scalars the semi-inner product can be recovered from the seminorm by

$$\begin{aligned}\langle G * \varphi, G * \psi \rangle_{w,\theta} &= \frac{1}{4}\left(|G * (\phi + \psi)|_{w,\theta}^2 - |G * (\phi - \psi)|_{w,\theta}^2\right) + \\ &\quad + \frac{i}{4}\left(|G * (\psi - i\phi)|_{w,\theta}^2 - |G * (\psi + i\phi)|_{w,\theta}^2\right) \\ &= \int \frac{f(\phi, \psi)}{w\left|\cdot\right|^{2\theta}},\end{aligned}$$

where

$$f(\phi, \psi) = \frac{1}{4}\left(\left|\widehat{\varphi} + \widehat{\psi}\right|^2 - \left|\widehat{\varphi} - \widehat{\psi}\right|^2\right) + \frac{i}{4}\left(\left|\widehat{\psi} - i\widehat{\phi}\right| - \left|\widehat{\psi} + i\widehat{\phi}\right|\right) = \widehat{\varphi}\overline{\widehat{\psi}},$$

and so the first equation of 6.1 holds.

Now if $\varphi, \psi \in \widehat{S}_{\emptyset,\theta}$ then by part 2 of Theorem 15, $\widehat{\varphi}\overline{\widehat{\psi}} \in S_{\emptyset,2\theta}$. Thus by Definition 44 of a basis distribution of order θ we have

$$\langle G * \varphi, G * \psi \rangle_{w,\theta} = \int \frac{\widehat{\varphi}\overline{\widehat{\psi}}}{w\left|\cdot\right|^{2\theta}} = \left[\widehat{G}, \widehat{\varphi}\overline{\widehat{\psi}}\right] = \left[G, \left(\widehat{\varphi}\overline{\widehat{\psi}}\right)^{\wedge}\right] = \left[G, \varphi_{\_} * \overline{\psi}\right],$$

where we have used the Fourier transform and convolution identities listed in the Appendix. ■

Definition 204 ***The spaces*** $\dot{J}_G$ ***and*** J_G. *Suppose the weight function w has property W2. Suppose G is a basis distribution of order $\theta \geq 0$ generated by a weight function w. Then the spaces $\dot{J}_G$ and J_G are defined by:*

$$\dot{J}_G = G * \widehat{S}_{\emptyset,\theta}, \qquad J_G = \dot{J}_G \oplus P_\theta.$$

Here $\widehat{S}_{\emptyset,\theta}$ denotes the Fourier transform of the functions in $S_{\emptyset,\theta}$ and $G * \widehat{S}_{\emptyset,\theta}$ denotes the convolution of the tempered distribution G with the functions in $\widehat{S}_{\emptyset,\theta}$.

The use of the direct product $\oplus$ in the definition of J_G must be justified and we must show that the definition is independent of the particular basis function chosen.

Theorem 205 *Suppose the weight function w has property W2. Suppose G is a basis distribution of order $\theta \geq 1$ generated by a weight function w. Then:*

1. $\dot{J}_G \cap P = \{0\}$.
2. $P_{2\theta} * \widehat{S}_{\emptyset,\theta} \subset P_\theta$.
3. $\dot{J}_{G_1} = \dot{J}_{G_2}$ *iff* $G_1 - G_2 \in P_\theta$.
4. *Set-wise, J_G is independent of the basis function G of order θ used to define it.*
5. *J_G is dense in X_w^θ when endowed with the Light norm.*

Proof. Part 1 Suppose $\dot{J}_G \cap P \neq \{0\}$. Then we can choose $p \in P$, $p \neq 0$ and $\varphi \in \widehat{S}_{\emptyset,\theta}$ such that $p = G * \varphi$. But $p \in X_w^\theta$ implies $p \in P_\theta$ so from equation 6.1

$$0 = |p|_{w,\theta}^2 = |G * \varphi|_{w,\theta}^2 = \int \frac{|\widehat{\varphi}|^2}{w\,|\cdot|^{2\theta}},$$

which implies $\widehat{\varphi} = 0$, and so $\varphi = 0$ and $p = 0$.

Part 2 Suppose $q \in P_{2\theta}$. Then by part 1 of Definition 44 $q = G_1 - G_2$ for some basis distributions G_1 and G_2. Thus by Theorem 203, $q * \widehat{S}_{\emptyset,\theta} = G_1 * \widehat{S}_{\emptyset,\theta} - G_2 * \widehat{S}_{\emptyset,\theta} \subset X_w^\theta$. But $q * \widehat{S}_{\emptyset,\theta}$ is a space of polynomials so $q * \widehat{S}_{\emptyset,\theta} \subset P_\theta$.

Part 3 Suppose $\dot{J}_{G_1} = \dot{J}_{G_2}$. Then given $\phi \in S_{\emptyset,\theta}$ there exists $\psi \in S_{\emptyset,\theta}$ such that $G_1 * \widehat{\phi} = G_2 * \widehat{\psi}$. Thus $G_2 * \left(\widehat{\psi} - \widehat{\phi}\right) = (G_1 - G_2) * \widehat{\phi} = q * \widehat{\phi} \in P$, for some $q \in P_{2\theta}$.

Part 1 of this theorem now implies that $\psi = \phi$ and $q * \widehat{\phi} = 0$. Indeed, we can conclude that $q * \widehat{\phi} = 0$ for all $\phi \in S_{\emptyset,\theta}$, or equivalently, $\phi\widehat{q} = 0$ for all $\phi \in S_{\emptyset,\theta}$. But part 2 of Theorem 17 implies that $q \in P_\theta$ and so $G_1 - G_2 \in P_\theta$.

Conversely, if $G_1 - G_2 \in P_\theta$ and $\phi \in S_{\emptyset,\theta}$ then $\phi\,(G_1 - G_2)^\vee = 0$ for all $\phi \in S_{\emptyset,\theta}$, or $(G_1 - G_2) * \phi = 0$ for all $\phi \in S_{\emptyset,\theta}$. Hence $\dot{J}_{G_1} = \dot{J}_{G_2}$.

Part 4 By Definition 44, $G_1 - G_2 \in P_{2\theta}$. Set $q = G_1 - G_2$. Now suppose $u \in J_{G_1}$, so that $u = G_1 * \phi + p$ for some $\phi \in \widehat{S}_{\emptyset,\theta}$ and $p \in P_\theta$. Consequently

$$u = G_1 * \phi + p = G_2 * \phi + (G_1 - G_2) * \phi + p = G_2 * \phi + q * \phi + p,$$

and by part 2 the last two terms are members of J_{G_2}. Repeating the argument for $u \in J_{G_1}$ proves this part.

Part 5 A standard result is that a subspace of a Hilbert space is dense iff its orthogonal complement is $\{0\}$.

Suppose $(f, G * \phi + p)_{w,\theta} = 0$ for all $\phi \in \widehat{S}_{\emptyset,\theta}$ and $p \in P_\theta$. Now when $p = -\mathcal{P}\,(G * \phi)$

$$\begin{aligned} 0 = (f, G * \phi + p)_{w,\theta} &= \langle f, G * \phi + p\rangle_{w,\theta} + \sum_k f\,(a_k)\,\overline{(G * \phi + p)\,(a_k)} \\ &= \langle f, G * \phi\rangle_{w,\theta} \\ &= \left[f, \overline{\phi}\right], \end{aligned}$$

where the last step used 6.2. Hence $\left[f, \overline{\phi}\right] = 0$ for all $\phi \in \widehat{S}_{\emptyset,\theta}$ and by Theorem 17 $f \in P_\theta$. Now we have $(f, p)_{w,\theta} = 0$ for all $p \in P_\theta$, which implies $\|f\|_{w,\theta}^2 = 0$ and so $f = 0$. ∎

6.3 Formulation of the Approximate smoothing problem

The Approximate smoother problem is derived from the Exact smoother problem studied in Chapter 5 by restricting the range of the minimizing functions from X_w^θ to a space $W_{G,X'}$, where X' is an arbitrary

set of θ-unisolvent points in $\mathbb{R}^d$. The first step is to assume that X' is a regular grid containing the Exact smoother data and to approximate or discretize the functions in X_w^θ using this grid. This leads to the finite dimensional subspace $W_{G,X'}$ and the Approximate smoother problem, namely $\min\limits_{f \in W_{G,X'}} J_e[f]$, where J_e is the Exact smoothing functional. We then generalize this problem to an arbitrary unisolvent X'.

In this section we will provide some justification for approximating the infinite dimensional Hilbert space X_w^θ by a finite dimensional subspace $W_{G,X'}$, where X' is a regular, rectangular grid of points in $\mathbb{R}^d$. The space $W_{G,X'}$ will be used to define the Approximate smoothing problem. The set X' will then be generalized to include any finite set of distinct points.

Definition 206 ***A regular, rectangular grid in*** $\mathbb{R}^d$

Let the grid occupy a rectangle $R(a;b)$, which has left-most point $a \in \mathbb{R}^d$ and right-most point b. Suppose the grid has $\mathcal{N}' = (N_1', N_2', \ldots, N_d')$ points in each dimension and let $h \in \mathbb{R}^d$ denote the grid sizes.

Then $X' = \left\{x_\alpha' = a + h\alpha \mid \alpha \in \mathbb{Z}^d \text{ and } 0 \leq \alpha < \mathcal{N}'\right\}$ is the set of grid points.

Let $N^{'}$ be the number of grid points so that $N^{'} = (\mathcal{N}')^{\mathbf{1}} = \mathcal{N}_1'\mathcal{N}_2' \ldots \mathcal{N}_d'$, and of course we have the constraint $\mathcal{N}^{'}h = b - a$.

The definition of the space $W_{G,X'}$ requires that X' is unisolvent. If X' is a regular, rectangular grid the next theorem shows that if the grid is made finer in all dimensions the grid eventually becomes unisolvent, no matter what order of unisolvency is required. We will need the following lemma:

Lemma 207 *We have the following unisolvency results:*

1. *The set $\left\{\gamma \in \mathbb{Z}^d : 0 \leq \gamma < n\right\}$ is unisolvent w.r.t. P_n.*
2. *Translations of minimal unisolvent sets are minimal unisolvent sets.*
3. *Dilations of minimal unisolvent sets are minimal unisolvent sets.*

Proof. Part 1. From the definition of unisolvency, Definition 99, we must show that for each $p \in P_n$, $p(\gamma) = 0$ for $0 \leq \gamma < n$ implies $p = 0$. The proof will be by induction on the order of the polynomial.

Clearly the lemma is true for $n = 1$ since P_1 is the constant polynomials.

Now assume that $n \geq 2$ and that if $p \in P_n$ and $p(\gamma) = 0$ for $0 \leq \gamma < n$ then $p = 0$. Set $p(x) = \sum\limits_{|\beta|<n} c_\beta x^\beta$. Then if $\gamma_k = 0$ and $\gamma_i = 1$ when $i \neq k$ then $0 = \sum\limits_{|\beta|<n} c_\beta \gamma^\beta$ implies $c_\beta = 0$ when $\beta_k = 0$. Thus $p(\gamma) = 0$ for $0 \leq \gamma < n$ implies $c_\beta = 0$ when $\beta_i = 0$ for some i. Consequently, $p = 0$ if $n \leq d$ else $p(x) = \sum\limits_{\substack{|\beta|<n \\ \beta>0}} c_\beta x^\beta$. If $n > d$ we can write $p(x) = x^{\mathbf{1}} q(x)$ where $q \in P_{n-1}$, so that $q(\gamma) = 0$ for $1 \leq \gamma < n$ must imply $q = 0$. Finally, if we define $r \in P_{n-1}$ by $r(x) = q(x + 1)$ then $r(\gamma) = 0$ for $0 \leq \gamma < n - 1$ must imply $r = 0$, and so the truth of our lemma for $n - 1$ implies the truth of the lemma for n, and the lemma is proved.

Parts 2 and 3. From the definition of a cardinal basis, Definition 100, there is a unique cardinal basis $\{l_i\}$ of P_n associated with a set $A = \{a_i\}$ iff A is minimally unisolvent of order n. Now by definition $l_i(a_j) = \delta_{i,j}$. Hence, if $\tau, \delta \in \mathbb{R}^d$ and δ has positive components, then the cardinal basis associated with the translation $A + \tau$ is $\{l_i(\cdot - \tau)\}$ and the cardinal basis associated with the dilation δA is $\{l_i(\cdot/\delta)\}$. ■

Theorem 208 *Suppose $X' = \left\{x_\alpha' = a + h\alpha \mid \alpha \in \mathbb{Z}^d \text{ and } 0 \leq \alpha < \mathcal{N}'\right\}$ is the regular, rectangular grid introduced in Definition 206. Then X' is θ-unisolvent if $\mathcal{N}' \geq \theta$.*

Proof. Since $(X' - a)/h = \left\{\alpha \mid \alpha \in \mathbb{Z}^d \text{ and } 0 \leq \alpha < \mathcal{N}'\right\}$ and $\mathcal{N}' \geq \theta$, part 1 of the lemma implies $(X' - a)/h$ is θ-unisolvent and thus from the definition of unisolvency, Definition 99, $(X' - a)/h$ must contain a minimal unisolvent subset. Parts 2 and 3 of the lemma imply that X' contains a minimal unisolvent subset and so X' is unisolvent. ■

By Theorem 205 the space $J_G = G * \widehat{S}_{\emptyset,\theta} + P_\theta$ is dense in X_w^θ under the Light norm sense. So we will approximate functions in $S_{\emptyset,\theta}$ using the grid X' defined above. Our analysis will be matrix-based so order the grid points and write $X^{'} = \{x_n'\}_{n=1}^{N^{'}}$. We will approximate integrals on the grid region using the trapezoidal rule i.e.

$$\int\limits_{grid} f(x)\,dx \simeq h^{\mathbf{1}} \sum_{n=1}^{N^{'}} f(x_n'), \tag{6.3}$$

where $h^{\mathbf{1}} = h_1 h_2 \times \ldots \times h_d$.

Step1 Approximation of the functions in $\dot{J}_G = G * \widehat{S}_{\emptyset,\theta}$ by functions in $\dot{W}_{G,X'}$.

We now need a basis for P_θ, say $\{q_l\}_{l=1}^{M}$ where $M = \dim P_\theta$. Then $\widehat{S}_{\emptyset,\theta}$ has the following characterization which we will not prove here:

$$\widehat{S}_{\emptyset,\theta} = \left\{ \phi \in S : \int q_l\left(x\right) \phi\left(x\right) dx = 0, \ for\ all\ l \right\}. \tag{6.4}$$

Suppose $\phi \in \widehat{S}_{\emptyset,\theta}$. Then the trapezoidal approximation 6.3 on the grid X' we have

$$G * \phi = \int\limits_{\mathbb{R}^d} G\left(x-y\right) \phi\left(y\right) dy \simeq \sum_{n=1}^{N'} G\left(x - x_n'\right) \left(h^1 \phi\left(x_n'\right)\right), \tag{6.5}$$

and

$$\int\limits_{\mathbb{R}^d} q_l\left(x\right) \phi\left(x\right) dx \simeq \sum_{n=1}^{N'} q_l\left(x_n'\right) \left(h^1 \phi\left(x_n'\right)\right). \tag{6.6}$$

Considering the equations and approximations 6.4, 6.5 and 6.6 we will approximate the space $G * \widehat{S}_{\emptyset,\theta}$ by functions of the form $\sum\limits_{n=1}^{N'} G\left(x - x_n'\right) \alpha_n$, $\alpha_n \in \mathbb{C}$ constrained by $\sum\limits_{n=1}^{N'} q_l\left(x_n'\right) \alpha_n = 0$ for $l = 1, \ldots, M$. In matrix terms these constraints become $P_{X'}^T \alpha = 0$, where $\alpha = (\alpha_n)$ and $P_{X'} = \left(q_j\left(x_i'\right)\right)$ is the unisolvent matrix introduced in Definition 103. But noting Definition 163 we see that our approximating space is just $\dot{W}_{G,X'}$, provided X' is unisolvent. But by Lemma 208 this is true if $N' > \dim P_\theta$.

Step 2 Approximate $J_G = G * \widehat{S}_{\emptyset,\theta} + P_\theta$ by $W_{G,X'} = \dot{W}_{G,X'} + P_\theta$.

Step 3 With this motivation we could now specify a smoothing problem, which we will call an Approximate smoothing problem, which involves minimizing the Exact smoothing functional J_e given by 5.6 over $W_{G,X'}$ where X' is a rectangular grid. However, since the space $W_{G,X'}$ is defined when X' is any θ-unisolvent set of distinct points we will define the following more general problem:

Definition 209 ***The Approximate smoothing problem***
Minimize the Exact smoothing functional $J_e\left[f\right]$ for $f \in W_{G,X'}$, where X' is a θ-unisolvent set of distinct points in $\mathbb{R}^d$. More concisely we can write $\min\limits_{f \in W_{G,X'}} J_e\left[f\right]$.

The Exact smoothing functional $J_e\left[f\right]$ is given by 5.6 and is defined using the scattered data $[X, y]$ where $X = \left\{x^{(i)}\right\}_{i=1}^{N}$ is the θ-unisolvent independent data and $y = \{y_i\}_{i=1}^{N}$ is termed the dependent data.

6.4 Solving the Approximate smoothing problem using Hilbert space techniques

In this section Hilbert space techniques are used to show the Approximate smoothing problem has a unique solution in the finite dimensional space $W_{G,X'}$. We then prove several identities satisfied by the smoothing function and obtain a matrix equation for the coefficients of the basis functions of the space $W_{G,X'}$.

6.4.1 Summary of Exact smoother properties

To start with we will require the following properties of the mapping $\mathcal{L}_X$ and the Exact smoother which were proved in Section 5.3:

Summary 210 *Suppose $[X, y]$ is the data for the* ***Exact smoothing problem*** *and X is a θ-unisolvent set. Assume the operators $\mathcal{P}$, $\mathcal{Q}$ and the Light norm $\left\|\cdot\right\|_{w,\theta}$ are all constructed using the same minimal unisolvent subset of X. Then:*

1. *If* $\zeta = (0, y)$ *then* $\|\mathcal{L}_X f - \varsigma\|_V^2 = J_e[f]$ *for* $f \in X_w^\theta$ *(Remark 169).*

   From Theorem 170:

2. $\mathcal{L}_X : X_w^\theta \to V$ *is continuous and 1-1.*

3. *If* $u = (u_1, \widetilde{u}_2) \in V$ *then* $\mathcal{L}_X^* u = \rho \mathcal{Q} u_1 + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{u}_2$ *w.r.t. the Light norm constructed from a minimal unisolvent subset of* X*.*

4. $\mathcal{L}_X^* \mathcal{L}_X f = \rho \mathcal{Q} f + \frac{1}{N} \widetilde{\mathcal{E}}_X^* \widetilde{\mathcal{E}}_X f$ *when* $f \in X_w^\theta$*. Also,* $\mathcal{L}_X^* \mathcal{L}_X : X_w^\theta \to X_w^\theta$ *is a homeomorphism and* $\mathcal{L}_X^* \mathcal{L}_X : W_{G,X} \to W_{G,X}$ *is a homeomorphism.*

   From Theorem 171:

5. $\|\mathcal{L}_X s_e - \varsigma\|_V^2 + \|\mathcal{L}_X s_e - \mathcal{L}_X f\|_V^2 = \|\mathcal{L}_X f - \varsigma\|_V^2$ *for all* $f \in X_w^\theta$*.*

6. *The Exact smoother problem has a unique solution* s_e *in* $W_{G,X}$ *and this preserves polynomials in* P_θ*.*

7. *The Exact smoother of the data* $[X, y]$ *is given by* $s_e = \frac{1}{N} \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} \widetilde{\mathcal{E}}_X^* y$*.*

6.4.2 The existence and uniqueness of the Approximate smoother

The geometric Hilbert space technique of orthogonal projection is used to show the Approximate smoothing problem has a unique solution in $W_{G,X'}$.

Theorem 211 ***The Approximate smoothing problem*** *of Definition 209 has a unique solution in* $W_{G,X'}$*, say* s_a*, which satisfies:*

1. $J_e[s_a] < J_e[f]$ *for all* $f \in W_{G,X'}$ *and* $f \neq s_a$*.*

2. $(\mathcal{L}_X s_a - \varsigma, \mathcal{L}_X s_a - \mathcal{L}_X f)_V = 0$ *for all* $f \in W_{G,X'}$*, where* $\varsigma = (0, y) \in V$*.*

3. $\|\mathcal{L}_X s_a - \varsigma\|_V^2 + \|\mathcal{L}_X s_a - \mathcal{L}_X f\|_V^2 = \|\mathcal{L}_X f - \varsigma\|_V^2$ *for all* $f \in W_{G,X'}$*. This is equivalent to part 2.*

4. $\left(\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y,\, f\right)_{w,\theta} = 0$ *for all* $f \in W_{G,X'}$*. This is equivalent to part 2.*

5. *The Approximate smoother is independent of the basis function* G *chosen to construct* $W_{G,X'}$*.*

6. *If* s_e *is the Exact smoother of the data* y *then* $s_a - s_e = \left(\mathcal{L}_X^* \mathcal{L}_X\right)^{-1} g$ *for some unique* $g \in W_{G,X'}^\perp$*.*

7. $\rho \left|s_e - s_a\right|_{w,\theta}^2 + \frac{1}{N} \sum\limits_{k=1}^{N} \left|s_e\left(x^{(k)}\right) - s_a\left(x^{(k)}\right)\right|^2 = J_e[s_a] - J_e[s_e]$*.*

Proof. Part 1 From part 1 of Summary 210 $J_e[f] = \|\mathcal{L}_X f - \varsigma\|_V^2$. So now we want to show that there is a unique function $f \in W_{G,X'}$ which minimizes the functional $\|\mathcal{L}_X f - \varsigma\|_V^2$ over $W_{G,X'}$. By part 3 Theorem 166, $W_{G,X'}$ is a finite dimensional subspace of X_w^θ so $\mathcal{L}_X(W_{G,X'})$ must be a finite dimensional subspace of V and hence a closed subspace of V. Consequently there exists a unique element of $\mathcal{L}_X(W_{G,X'})$, say v, which is the orthogonal projection of ζ onto $\mathcal{L}_X(W_{G,X'})$ such that $\|v - \varsigma\|_V < \|\mathcal{L}_X f - \varsigma\|_V$ for all $f \in W_{G,X'}$ and $\mathcal{L}_X(f) \neq v$.

Since $\mathcal{L}_X$ is 1-1 on X_w^θ there exists a unique element of $W_{G,X'}$, call it s_a, such that $v = \mathcal{L}_X(s_a)$.

In terms of J_e we have $J_e[s_a] < J_e[f]$ for all $f \in W_{G,X'}$ and $f \neq s_a$.

Parts 2 and **3** Since v is the projection of ζ onto $\mathcal{L}_X(W_{G,X'})$ we have the equivalent equations of parts 2 and 3.

Part 4 We study the equation of part 2 using the properties of the operators $\mathcal{L}_X f = \left(f, \widetilde{\mathcal{E}}_X f\right) \in V$ and $\mathcal{L}_X^*$ given in Theorem 210. In fact, for all $g \in W_{G,X'}$

$$\begin{aligned} 0 = (\mathcal{L}_X s_a - \varsigma, \mathcal{L}_X s_a - \mathcal{L}_X g)_V = (\mathcal{L}_X s_a - \varsigma, \mathcal{L}_X(s_a - g))_V &= (\mathcal{L}_X^* \mathcal{L}_X s_a - \mathcal{L}_X^* \varsigma, s_a - g)_{w,\theta} \\ &= \left(\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y, s_a - g\right)_{w,\theta}. \end{aligned}$$

But $s_a \in W_{G,X'}$ so

$$\left(\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y, f \right)_{w,\theta} = 0, \quad f \in W_{G,X'},$$

and this is clearly equivalent to part 2.

Part 5 From part 1 of Definition 163 we know that the set $W_{G,X'}$ is independent of the basis function used in its construction, and since the Exact smoother functional 5.6 is independent of the basis function the result follows.

Part 6 Part 4 implies directly that $\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y \in W_{G,X'}^{\perp}$, say $\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y = g$, and this part follows since by part 7 of Summary 210, $s_e = \frac{1}{N} \left(\mathcal{L}_X^* \mathcal{L}_X \right)^{-1} \widetilde{\mathcal{E}}_X^* y$.

Part 7 Substitute $f = s_a$ in the identity 5.9:

$$J_e[s_e] + \rho \left| s_e - f \right|_{w,\theta}^2 + \frac{1}{N} \sum_{k=1}^{N} \left| s_e \left(x^{(k)} \right) - f \left(x^{(k)} \right) \right|^2 = J_e[f].$$

■

Remark 212 *In part 4 above we proved that* $\mathcal{L}_X^* \mathcal{L}_X s_a - \frac{1}{N} \widetilde{\mathcal{E}}_X^* y \in W_{G,X'}^{\perp}$. *It was necessary to assume that X was θ-unisolvent and that X' was θ-unisolvent. However, the minimal unisolvent subsets of X and X' used to construct $\widetilde{\mathcal{E}}_X^*$, $\mathcal{P}$, $W_{G,X'}$ etc. are not required to have any points in common. We must be careful to avoid any calculations which are dependent on both X and X' but which may only be valid when it is assumed that X and X' share a minimal unisolvent subset and this set is used to calculate $\mathcal{P}$, $\mathcal{Q}$, the Light norm etc. For example, consider the equation $\widetilde{\mathcal{E}}_{X'} \widetilde{\mathcal{E}}_X^* \alpha = R_{X',X} \alpha$. What minimal unisolvent subset is used to calculate R_x? In this example a minimal unisolvent subset of X is used to construct R_x which is then evaluated using X' and X.*

6.4.3 Various identities

The identities of this subsection are similar to the Exact smoother identities of Corollary 172 and Corollary 173. They relate the Hilbert space properties and the pointwise properties of the data and the Approximate smoother.

Corollary 213 *Suppose s_a is the Approximate smoother of the data $X = \left\{ x^{(i)} \right\}$ and $y = \{y_i\}$ induced by the points X', and $f_d \in X_w^\theta$ is a data function for the Exact smoother. Then for all $f \in W_{G,X'}$:*

1.

$$\rho \left| s_a \right|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^{N} \left| s_a \left(x^{(i)} \right) - y_i \right|^2 + \rho \left| s_a - f \right|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^{N} \left| s_a \left(x^{(i)} \right) - f \left(x^{(i)} \right) \right|^2$$
$$= \rho \left| f \right|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^{N} \left| f \left(x^{(i)} \right) - y_i \right|^2.$$

2. If $f_d \in W_{G,X'}$ then

$$\left| s_a \right|_{w,\theta}^2 + \frac{2}{N\rho} \sum_{i=1}^{N} \left| s_a \left(x^{(i)} \right) - f_d \left(x^{(i)} \right) \right|^2 + \left| s_a - f_d \right|_{w,\theta}^2 = \left| f_d \right|_{w,\theta}^2.$$

Proof. Part 1 Expand the equation of part 3 of Theorem 211 using the formula $\left\| \mathcal{L}_X u \right\|_V^2 = \rho \left| u \right|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^{N} \left| u \left(x^{(i)} \right) \right|^2$ implied by the definition of $\mathcal{L}_X$ and V (Definition 168).

Part 2 Set $f = f_d$ in part 1. ■

The result of part 4 of Theorem 211 is used to prove the next corollary.

Corollary 214 *Suppose $s_a \in W_{G,X'}$ is the (unique) Approximate smoother of the data $[X, y]$ induced by the points X'. Then:*

1. $\langle s_a, f\rangle_{w,\theta} = \frac{1}{N\rho}\left(y - (s_a)_X\right)^T \overline{f_X}, \quad f \in W_{G,X'}$.

2. $P_X^T\left(y - (s_a)_X\right) = 0$ *where* P_X *is any unisolvency matrix defined using* X.

3. $|s|_{w,\theta}^2 = \frac{1}{N\rho}\left(y - (s_a)_X\right)^T \left(\overline{s_a}\right)_X^T$.

4. $J_e(s_a) = \frac{1}{N}\left(y - (s_a)_X\right)^T \overline{y}$.

5. $\left|(s_a)_X\right|^2 \le (s_a)_X^T\,\overline{y} = \left(\overline{s_a}\right)_X^T y \le |y|^2$.

Proof. Part 1 By part 4 Summary 210, $\mathcal{L}_X^*\mathcal{L}_X s_a = \rho\mathcal{Q}s_a + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X s_a$ and by part 1 Theorem 116, $(\mathcal{Q}s_a, f)_{w,\theta} = \langle s_a, f\rangle_{w,\theta}$ so that if $f \in W_{G,X'}$,

$$
\begin{aligned}
0 = \left(\mathcal{L}_X^*\mathcal{L}_X s_a - \frac{1}{N}\widetilde{\mathcal{E}}_X^* y, f\right)_{w,\theta} &= \left(\rho\mathcal{Q}s_a + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\widetilde{\mathcal{E}}_X s_a - \frac{1}{N}\widetilde{\mathcal{E}}_X^* y, f\right)_{w,\theta} \\
&= \left(\rho\mathcal{Q}s_a + \frac{1}{N}\widetilde{\mathcal{E}}_X^*\left(\widetilde{\mathcal{E}}_X s_a - y\right), f\right)_{w,\theta} \\
&= \rho\langle s_a, f\rangle_{w,\theta} + \frac{1}{N}\left(\widetilde{\mathcal{E}}_X s_a - y, \widetilde{\mathcal{E}}_X f\right)_{\mathbb{C}^N} \qquad (6.7) \\
&= \rho\langle s_a, f\rangle_{w,\theta} + \frac{1}{N}\left(s - y, f_X\right)_{\mathbb{C}^N} \\
&= \rho\langle s_a, f\rangle_{w,\theta} + \frac{1}{N}\left((s_a)_X - y\right)^T \overline{f_X},
\end{aligned}
$$

and hence

$$
\langle s_a, f\rangle_{w,\theta} = \frac{1}{N\rho}\left(y - (s_a)_X\right)^T \overline{f_X}.
$$

Part 2 Suppose $\{p_i\}$ is a basis for P_θ. Then if $f = p_i \in P_\theta$ the equation proved in part 1 becomes $(p_i)_X^T\left(y - (s_a)_X\right) = 0$ and so $(P_X)^T\left((s_a)_X - y\right) = 0$ since $P_X = \left(p_j\left(x^{(i)}\right)\right)$.

Part 3 Let $f = s_a$ in the equation proved in part 1.

Part 4 By part 3, $|s_a|_{w,\theta}^2 = \frac{1}{N\rho}\left(y - (s_a)_X\right)^T\left(\overline{s_a}\right)_X^T$ so

$$
\begin{aligned}
J_e(s_a) &= \rho|s_a|_{w,\theta}^2 + \frac{1}{N}\left|(s_a)_X - y\right|^2 \\
&= \rho|s_a|_{w,\theta}^2 + \frac{1}{N}\left((s_a)_X - y^T\right)\left(\left(\overline{s_a}\right)_X^T - \overline{y}\right) \\
&= \rho|s_a|_{w,\theta}^2 + \frac{1}{N}\left((s_a)_X^T\left(\overline{s_a}\right)_X^T - (s_a)_X^T\,\overline{y} - y^T\left(\overline{s_a}\right)_X^T + y^T\overline{y}\right) \\
&= \frac{1}{N}\left(y - (s_a)_X\right)^T\left(\overline{s_a}\right)_X^T + \frac{1}{N}\left((s_a)_X^T\left(\overline{s_a}\right)_X^T - (s_a)_X^T\,\overline{y} - y^T\left(\overline{s_a}\right)_X^T + y^T\overline{y}\right) \\
&= \frac{1}{N}\left(y^T\left(\overline{s_a}\right)_X^T - (s_a)_X^T\left(\overline{s_a}\right)_X^T + (s_a)_X^T\left(\overline{s_a}\right)_X^T - (s_a)_X^T\,\overline{y} - y^T\left(\overline{s_a}\right)_X^T + y^T\overline{y}\right) \\
&= \frac{1}{N}\left(-(s_a)_X^T\,\overline{y} + y^T\overline{y}\right) \\
&= \frac{1}{N}\left(y - (s_a)_X\right)^T\overline{y}.
\end{aligned}
$$

Part 5 Part 3 implies $\left|(s_a)_X\right|^2 \le (s_a)_X^T\,\overline{y} = \left(\overline{s_a}\right)_X^T y$ and part 4 implies $(s_a)_X^T\,\overline{y} \le |y|^2$. ■

6.4.4 Matrices and vectors derived from the Riesz representer and the basis function

The interpolation and Exact smoother problems only involve a single independent data set X and this leads to matrix equations which only use matrices of the form $R_{X,X} = \left(R_{x^{(j)}}\left(x^{(i)}\right)\right)$ and $G_{X,.X} = \left(G\left(x^{(i)} - x^{(j)}\right)\right)$. However, the Approximate smoother problem (Definition 209) involves two independent data sets X and X' which will require the following definitions:

Definition 215 ***Matrices and vectors derived from the Riesz representer*** R_x ***and the basis function*** G
Suppose $Y = \{y^{(k)}\}$ *and* $Z = \{z^{(k)}\}$ *are arbitrary sets of points in* $\mathbb{R}^d$ *and* $y, z \in \mathbb{R}^d$. *Then:*

1. $R_{Y,Z} = \left(R_{z^{(j)}}\left(y^{(i)}\right)\right)$ *and* $G_{Y,Z} = \left(G\left(y^{(i)} - z^{(j)}\right)\right)$.
2. $R_{Y,z} = \left(R_z\left(y^{(i)}\right)\right)$ *and* $G_{Y,z} = \left(G\left(y^{(i)} - z\right)\right)$.
3. $R_{y,Z} = \left(R_{z^{(j)}}(y)\right)$ *and* $G_{y,Z} = \left(G\left(y - z^{(j)}\right)\right)$.
   $R_{Y,Z}$ *is called a (asymmetric)* ***reproducing kernel matrix*** *and* $G_{Y,Z}$ *is called a (asymmetric)* ***basis function matrix***.

The next result derives an important relationship between the reproducing kernel matrix and the basis function matrix.

Theorem 216 *Suppose now that* A *is a minimal unisolvent set with cardinal basis by* $\{l_i\}_{i=1}^M$. *Define the Riesz representer* R_x *using* A *and* $\{l_i\}_{i=1}^M$. *Then if* $Y = \{y^{(k)}\}$ *and* $Z = \{z^{(k)}\}$ *are arbitrary sets of points in* $\mathbb{R}^d$*:*

1. $$R_{Y,Z} = (2\pi)^{-\frac{d}{2}}\left(G_{Y,Z} - L_Y G_{A,Z} - G_{Y,A} L_Z^T + L_Y G_{A,A} L_Z^T\right) + L_Y L_Z^T. \tag{6.8}$$
2. *If* Y *and* Z *are sets of distinct points in* $\mathbb{R}^d$ *then* $\widetilde{\mathcal{E}}_Y \widetilde{\mathcal{E}}_Z^* = R_{Y,Z}$.

Proof. Part 1 From 4.35

$$(2\pi)^{\frac{d}{2}} R_z(y) = G(y - z) - \sum_{j=1}^{M} l_j(y) G(a_j - z) - \sum_{i=1}^{M} G(y - a_i) l_i(z) + \\ + \sum_{i,j=1}^{M} l_j(y) G(a_j - a_i) l_i(z) + (2\pi)^{\frac{d}{2}} \sum_{j=1}^{M} l_j(z) l_j(y),$$

or in the notation introduced in Definition 215

$$R_z(y) = (2\pi)^{-\frac{d}{2}}\left(G(y - z) - \widetilde{l}(y)^T G_{A,z} - G_{y,A}\widetilde{l}(z) + \widetilde{l}(y)^T G_{A,A}\widetilde{l}(z)\right) + \widetilde{l}(y)^T \widetilde{l}(z),$$

Now $R_{y,X}$ is the row vector $\left(R_{x^{(j)}}(y)\right)$ and $L_X = \left(l_j\left(x^{(i)}\right)\right)$ so

$$R_{y,Z} = (2\pi)^{-\frac{d}{2}}\left(G_{y,Z} - \widetilde{l}(y)^T G_{A,Z} - G_{y,A} L_Z^T + \widetilde{l}(y)^T G_{A,A} L_Z^T\right) + \widetilde{l}(y)^T L_Z^T,$$

and hence

$$R_{Y,Z} = (2\pi)^{-\frac{d}{2}}\left(G_{Y,Z} - L_Y G_{A,X} - G_{Y,A} L_Z^T + L_Y G_{X,X} L_Z^T\right) + L_Y L_Z^T$$

Part 2 Suppose $Y = \{y^{(i)}\}$, $Z = \{z^{(j)}\}$ and $\beta = (\beta_j)$. Then from Definition 138 and 4.50,

$$\widetilde{\mathcal{E}}_Y \widetilde{\mathcal{E}}_Z^* \beta = \sum_{j=1}^{N} \beta_j \widetilde{\mathcal{E}}_Y R_{z^{(j)}} = \sum_{j=1}^{N} \beta_j \left(R_{z^{(j)}}\left(y^{(i)}\right)\right) = R_{Y,Z}\beta,$$

where the last step used the definition of $R_{Y,Z}$ from Definition 215. ■

6.4.5 The Approximate smoother matrix equation

We now know the Approximate smoother exists, is unique and is a member of $W_{G,X'}$. The next step is to derive a matrix equation for the coefficients of the data-translated basis functions and the basis polynomials. This proof makes good use of the identities of Subsection 6.4.3 and the properties of unisolvency matrices, and is quite similar to the proof of Theorem 183 which derives the Exact smoother matrix equation. This proof will require notation to deal with inner products, Riesz representers etc. which are generated by two minimal unisolvent sets A and A':

Notation 217 *Suppose A and A' are two minimal unisolvent sets and A is the "default" notation. Let $(\cdot,\cdot)'_{w,\theta}$ be the Light inner product generated by A', R'_x be the Riesz representer generated by A' and $\widetilde{\mathcal{E}}^*_{A';X}$ be the adjoint of $\widetilde{\mathcal{E}}_X$ w.r.t. $(\cdot,\cdot)'_{w,\theta}$.*

Regarding unisolvency matrix notation, if the P_θ basis used for $P_{X'}$ is different to that used for P_X then use the notation $P'_{X'}$ instead of $P_{X'}$. In contrast, no 'prime' notation is required for the cardinal unisolvency matrix $L_{X'}$ because the (cardinal) basis used to calculate $L_{X'}$ is uniquely determined by A'.

Theorem 218 ***The Approximate smoother matrix equation***

Fix a basis function G of order θ. Choose a minimal unisolvent set $A' \subset X'$ and a P_θ basis $\{p_i\}$ to define the unisolvency matrix $P_{X'}$, the Riesz representer R'_x and the space $W_{G,X'}$. Choose a minimal unisolvent set $A \subset X$ and the P_θ basis $\{p_i\}$ to define the unisolvency matrix P_X, the Riesz representer R_x and $W_{G,X}$. Now suppose $s \in W_{G,X'}$ is the (unique) Approximate smoother of the data $\left[X = \left\{x^{(i)}\right\}_{i=1}^{N}, y = (y_i)\right]$ induced by the set of points $X' = \{x'_i\}_{i=1}^{N'}$. Since $s \in W_{G,X'}$

$$s(x) = \sum_{i=1}^{N'} \alpha'_i G(x - x'_i) + \sum_{i=1}^{M} \beta'_i p_i(x), \tag{6.9}$$

*for some $\alpha = (\alpha'_i) \in \mathbb{C}^N$ and $\beta' = (\beta'_i) \in \mathbb{C}^M$. In fact, α' and β' satisfy the **Approximate smoother matrix equation***

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X} G_{X,X'} & G_{X',X} P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix} \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix} = \begin{pmatrix} G_{X',X} \\ P_X^T \\ O_{M,N} \end{pmatrix} y, \tag{6.10}$$

where $G_{X',X'} = G\left(x'_i - x'_j\right)$, $G_{X',X} = G\left(x'_i - x^{(j)}\right)$, $G_{X',X} = G\left(x^{(i)} - x'_j\right)$ are basis function matrices and $P_X = \left(p_j\left(x^{(i)}\right)\right)$ and $P_{X'} = \left(p_j\left(x'_i\right)\right)$ are unisolvency matrices (Definition 103).

*The matrix on the left of this equation will be called the **Approximate smoother matrix** and will usually be denoted by the symbol Ψ.*

Proof. From part 3 of Definition 163 the space $W_{G,X'}$ is independent of the ordering of the points in X'. Thus the Approximate smoothing problem is also independent of the ordering of the points in X' and we now take advantage of this to order X' so that the first M points of X' lie in A'. Equation 6.7 of the proof of Corollary 214 is

$$\langle s, f\rangle_{w,\theta} = \frac{1}{N\rho}\left(y - \widetilde{\mathcal{E}}_X s, \widetilde{\mathcal{E}}_X f\right), \quad f \in W_{G,X'},$$

so the Light norm 4.29 corresponding to A' is

$$\begin{aligned} (s,f)'_{w,\theta} &= \frac{1}{N\rho}\left(y - \widetilde{\mathcal{E}}_X s, \widetilde{\mathcal{E}}_X f\right) + \left(\widetilde{\mathcal{E}}_{A'} s, \widetilde{\mathcal{E}}_{A'} f\right) \\ &= \frac{1}{N\rho}\left(\widetilde{\mathcal{E}}^*_{A';X}\left(y - \widetilde{\mathcal{E}}_X s\right), f\right)'_{w,\theta} + \left(\widetilde{\mathcal{E}}^*_{A';A'} \widetilde{\mathcal{E}}_{A'} s, f\right)'_{w,\theta}, \end{aligned}$$

where we have used Notation 217. By part 6 Theorem 166 the Riesz representer $R'_x \in W_{G,X'}$ when $x \in X'$ so the last equation implies

$$\widetilde{\mathcal{E}}_{X'}\left(s - \widetilde{\mathcal{E}}^*_{A';A'}\widetilde{\mathcal{E}}_{A'} s + \frac{1}{N\rho}\widetilde{\mathcal{E}}^*_{A';X}\left(s - \widetilde{\mathcal{E}}_X y\right)\right) = 0,$$

or

$$N\rho\widetilde{\mathcal{E}}_{X'} s - N\rho\widetilde{\mathcal{E}}_{X'}\widetilde{\mathcal{E}}^*_{A';A'}\widetilde{\mathcal{E}}_{A'} s + \widetilde{\mathcal{E}}_{X'}\widetilde{\mathcal{E}}^*_{A';X}\widetilde{\mathcal{E}}_X s - \widetilde{\mathcal{E}}_{X'}\widetilde{\mathcal{E}}^*_{A';X} y = 0.$$

By part 5 Theorem 139 and Definition 215, $\widetilde{\mathcal{E}}_{X'}\widetilde{\mathcal{E}}^*_{A';A'} = L_{X'}$ and $\widetilde{\mathcal{E}}_{X'}\widetilde{\mathcal{E}}^*_{A';X} = R'_{X';X}$ so

$$N\rho\widetilde{\mathcal{E}}_{X'} s - N\rho L_{X'}\mathcal{E}_{A'} s + R'_{X';X}\widetilde{\mathcal{E}}_X s - R'_{X';X} y = 0,$$

and for clarity we revert to the subscript notation for evaluations:

$$N\rho s_{X'} - N\rho L_{X'} s_{A'} + R'_{X';X} s_X - R'_{X';X} y = 0,$$

or

$$N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+R'_{X',X}\left(s_X-y\right)=0 \tag{6.11}$$

where the augmented square matrix $L_{X':0}=\begin{pmatrix}L_{X'} & O_{N,N-M}\end{pmatrix}$ was introduced in part 3 Theorem 105. Setting $Y=X'$ and $Z=X$ in formula 6.8 and using the special ordering of X' mentioned at the start of the proof we obtain

$$(2\pi)^{\frac{d}{2}}R'_{X',X}=G_{X',X}-L_{X'}G_{A',X}-G_{X',A'}L_X^T+L_{X'}G_{A',A'}L_X^T+(2\pi)^{\frac{d}{2}}L_{X'}L_X^T. \tag{6.12}$$

Part 2 of Corollary 214 implies $L_X^T\left(s_X-y\right)=0$. Using this equation and 6.12, 6.11 becomes

$$\begin{aligned}
0 &= (2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+(2\pi)^{\frac{d}{2}}R'_{X',X}\left(s_X-y\right)\\
&= (2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+\left(G_{X',X}-L_{X'}G_{A',X}-G_{X',A'}L_X^T+L_{X'}G_{A',A'}L_X^T\right)\left(s_X-y\right)+\\
&\qquad +(2\pi)^{\frac{d}{2}}L_{X'}L_X^T\left(s_X-y\right)\\
&= (2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+\left(G_{X',X}-L_{X'}G_{A',X}\right)\left(s_X-y\right)\\
&= (2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+\left(G_{X',X}-L_{X':0}G_{X',X}\right)s_X-\left(G_{X',X}-L_{X':0}G_{X',X}\right)y\\
&= (2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+\left(I_{N'}-L_{X':0}\right)G_{X',X}s_X-\left(I_{N'}-L_{X':0}\right)G_{X',X}y.
\end{aligned} \tag{6.13}$$

The next step is to express the factors $s_{X'}$ and s_X in basis function terms. But from the statement of this theorem, $s(x)=\sum\limits_{i=1}^{N'}\alpha'_iG\left(x-x'_i\right)+\sum\limits_{j=1}^{M}\beta'_jp_j(x)$ constrained by

$$P_{X'}^T\alpha'=0, \tag{6.14}$$

so that

$$s_X=G_{X,X'}\alpha'+P_X\beta',\quad s_{X'}=G_{X',X'}\alpha'+P_{X'}\beta'. \tag{6.15}$$

From part 3 of Theorem 104 we know that $P_{X'}=L_{X'}P_{A'}$, where $P_{A'}$ is regular. Also from part 3 Theorem 105 $L_{X':0}L_{X'}=L_{X'}$. Hence

$$\begin{aligned}
&(2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)s_{X'}+\left(I_{N'}-L_{X':0}\right)G_{X',X}s_X\\
&\quad=(2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)\left(G_{X',X'}\alpha'+P_{X'}\beta'\right)+\left(I_{N'}-L_{X':0}\right)G_{X',X}s_X\\
&\quad=(2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)\left(G_{X',X'}\alpha'+L_{X'}P_{A'}\beta'\right)+\left(I_{N'}-L_{X':0}\right)G_{X',X}s_X\\
&\quad=(2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)G_{X',X'}\alpha'+\left(I_{N'}-L_{X':0}\right)G_{X',X}s_X\\
&\quad=(2\pi)^{\frac{d}{2}}N\rho\left(I_{N'}-L_{X':0}\right)G_{X',X'}\alpha'+\left(I_{N'}-L_{X':0}\right)G_{X',X}\left(G_{X,X'}\alpha'+P_X\beta'\right)\\
&\quad=\left(I_{N'}-L_{X':0}\right)\left(\left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}+G_{X',X}G_{X,X'}\right)\alpha'+G_{X',X}P_X\beta'\right)\\
&\quad=\left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}+G_{X',X}G_{X,X'}\right)\alpha'+G_{X',X}P_X\beta'-\\
&\quad-L_{X':0}\left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}\alpha'+G_{X',X}G_{X,X'}\alpha'+G_{X',X}P_X\beta'\right),
\end{aligned}$$

and 6.13 becomes

$$\begin{aligned}
0 &= \left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}+G_{X',X}G_{X,X'}\right)\alpha'+G_{X',X}P_X\beta'-\\
&\qquad -L_{X':0}\left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}\alpha'+G_{X',X}G_{X,X'}\alpha'+G_{X',X}P_X\beta'\right)\\
&\qquad -\left(I_{N'}-L_{X':0}\right)G_{X',X}y\\
&= \left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}+G_{X',X}G_{X,X'}\right)\alpha'+G_{X',X}P_X\beta'-G_{X',X}y-\\
&\qquad -L_{X':0}\left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}\alpha'+G_{X',X}G_{X,X'}\alpha'+G_{X',X}P_X\beta'+G_{X',X}y\right)\\
&= \left((2\pi)^{\frac{d}{2}}N\rho G_{X',X'}+G_{X',X}G_{X,X'}\right)\alpha'+G_{X',X}P_X\beta'-G_{X',X}y-\\
&\qquad -L_{X'}\left((2\pi)^{\frac{d}{2}}N\rho G_{A',X'}\alpha'+G_{A',X}G_{X,X'}\alpha'+G_{A',X}P_X\beta'+G_{A',X}y\right),
\end{aligned}$$

From part 3 of Theorem 104 we know that $L_{X'} = P_{X'}P_{A'}^{-1}$, where $P_{A'}$ is regular, so

$$\begin{aligned}0 = &\left((2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'}\right)\alpha' + G_{X',X}P_X\beta' - G_{X',X}y-\\ &- P_{X'}P_{A'}^{-1}\left((2\pi)^{\frac{d}{2}} N\rho G_{A',X'}\alpha' + G_{A',X}G_{X,X'}\alpha' + G_{A',X}P_X\beta' + G_{A',X}y\right),\end{aligned}$$

which we write as

$$\left((2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'}\right)\alpha' + G_{X',X}P_X\beta' + P_{X'}\gamma' = G_{X',X}y, \tag{6.16}$$

where

$$\gamma' = P_{A'}^{-1}\left((2\pi)^{\frac{d}{2}} N\rho G_{A',X'}\alpha' + G_{A',X}G_{X,X'}\alpha' + G_{A',X}P_X\beta' + G_{A',X}y\right).$$

Finally, from part 2 of Corollary 214 it follows that $P_X^T(s_X - y) = 0$ and thus substituting for s_X from 6.15 we obtain

$$P_X^T G_{X,X'}\alpha' + P_X^T P_X\beta' = P_X^T y. \tag{6.17}$$

Equations 6.14, 6.16 and 6.17 now combine to give the Approximate smoother (block) matrix equation 6.10.

At the start of this proof we assumed a special order for the points in the ordered set X'. To prove that the Approximate smoother is unchanged by reordering the points X' we use the permutation matrix approach. In fact we will show that for the permutation π of X'

$$\begin{pmatrix} \begin{matrix}(2\pi)^{\frac{d}{2}} N\rho G_{\pi(X'),\pi(X')}+\\ +G_{\pi(X'),X}G_{X,\pi(X')}\end{matrix} & G_{\pi(X'),X}P_X & P_{\pi(X')} \\ P_X^T G_{X,\pi(X')} & P_X^T P_X & O_M \\ P_{\pi(X')}^T & O_M & O_M \end{pmatrix}\begin{pmatrix}\pi(\alpha') \\ \beta' \\ \gamma'\end{pmatrix} = \begin{pmatrix} G_{\pi(X'),X} \\ P_X^T \\ O_{M,N}\end{pmatrix} y, \tag{6.18}$$

and

$$s(x) = G_{x,\pi(X')}\pi(\alpha') + \sum_{i=1}^{M}\beta_i p_i(x). \tag{6.19}$$

Denote the permutation matrix of π by Π. To reorder the rows of a matrix we left-multiply by Π and to reorder the columns we right-multiply by Π^T. Also $\Pi^T\Pi = \Pi\Pi^T = I$. Hence, $G_{\pi(X'),\pi(X')} = \Pi G_{X',X'}\Pi^T$, $G_{X,\pi(X')} = G_{X,X'}\Pi^T$, $G_{\pi(X'),X} = \Pi G_{X',X}$ and $P_{\pi(X')} = \Pi P_{X'}$. The left side of the equation of the first row of 6.18 now becomes

$$\begin{aligned}&\left((2\pi)^{\frac{d}{2}} N\rho G_{\pi(X'),\pi(X')} + G_{\pi(X'),X}G_{X,\pi(X')}\right)\pi(\alpha') + G_{\pi(X'),X}P_X\beta' + P_{\pi(X')}\gamma'\\ &= \left((2\pi)^{\frac{d}{2}} N\rho\Pi G_{X',X'}\Pi^T + \Pi G_{X',X}G_{X,X'}\Pi^T\right)\Pi\alpha' + \Pi G_{X',X}P_X\beta' + \Pi P_{X'}\gamma'\\ &= \Pi\left(\left((2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'}\right)\Pi^T\Pi\alpha' + G_{X',X}P_X\beta' + P_{X'}\gamma'\right)\\ &= \Pi\left(\left((2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'}\right)\alpha' + G_{X',X}P_X\beta' + P_{X'}\gamma'\right)\\ &= \Pi G_{X',X}y\\ &= G_{\pi(X'),X},\end{aligned}$$

where the last line is implied by 6.10. Thus the equation of the first row of 6.18 is true, and the other equations are proved in a similar manner.

Using the notation of Definition 215 we can write 6.9 as

$$s(x) = G_{x,X'}\alpha' + \sum_{i=1}^{M}\beta_i' p_i(x) = G_{x,X'}\Pi^T\Pi\alpha' + \sum_{i=1}^{M}\beta_i' p_i(x) = G_{X,\pi(X')}\pi(\alpha') + \sum_{i=1}^{M}\beta_i' p_i(x),$$

which proves 6.19 and finishes the proof. ■

Remark 219 *A different basis for P_θ, say $\{p_j'\}$, could be used to define the unisolvency matrix $P_{X'}$, the Riesz representer R_x' and the space $W_{G,X'}$. Using the conventions of Notation 217 we would write $P_{X'}'$ instead of $P_{X'}$. All the results proved below would still hold.*

Recall from Definition 44 that the set of basis functions of order θ is $G + P_{2\theta}$, where G is a particular basis function. However, from part 5 of Theorem 211 we know that the Approximate smoother function is independent of the basis function chosen. The next corollary tells us that we can always choose a basis function such that the Approximate smoother matrix is Hermitian and indeed this particular matrix allows a simple proof of regularity which is given in Theorem 221.

Corollary 220 *Suppose the weight function w has properties W2 and W3 for parameter θ.*

Then a (possibly complex-valued) basis function G of order θ can always be chosen so that $G_{Y,Z}^T = \overline{G_{Z,Y}}$ for all finite subsets Y and Z of $\mathbb{R}^d$ i.e. so that $G(-x) = \overline{G(x)}$ for $x \in \mathbb{R}^d$. For such a basis function the Approximate smoother matrix Ψ of 6.10 is Hermitian and we can write

$$\Psi = \begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + \overline{G_{X,X'}^T} G_{X,X'} & \overline{G_{X,X'}^T} P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix}, \tag{6.20}$$

where $M = \dim P_\theta$.

Proof. From part 2 Theorem 92 there exists a basis function G which satisfies $G(-x) = \overline{G(x)}$ for $x \in \mathbb{R}^d$. The condition $G_{Y,Z}^T = \overline{G_{Z,Y}}$ implies that $G_{X',X'}^T = G_{X',X'}$ and $G_{X,X'}^T = G_{X',X}$ and since P_X and $P_{X'}$ are real valued it follows that Ψ has the required Hermitian form 6.20. ■

In the next theorem we will prove the Approximate smoother matrix 6.20 is regular.

Theorem 221 *The Approximate smoother matrix Ψ specified in Corollary 220 has the following properties:*

1. *Ψ is a regular Hermitian matrix.*
2. *Ψ is square with $N' + 2M$ rows. Hence the size of Ψ is independent of the number N of (scattered) data points.*

Proof. Part 1 We first write Ψ in the form.

$$\Psi = \begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X';0} + \overline{G_P^T} G_P & \begin{matrix} P_{X'} \\ O_{M,N'} \end{matrix} \\ \begin{matrix} P_{X'}^T & O_{N',M} \end{matrix} & O_M \end{pmatrix},$$

where

$$G_{X';0} = \begin{pmatrix} G_{X',X'} & O \\ O & O_M \end{pmatrix}, \quad G_P = \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix}.$$

Since $G_{X,X}$ is Hermitian Ψ is Hermitian. To prove that Ψ is regular we use Lemma 143. To this end set

$$\begin{aligned} B &= (2\pi)^{\frac{d}{2}} N\rho G_{X';0} + \overline{G_P^T} G_P, \\ C &= \begin{pmatrix} P_{X'} \\ O_{M,N'} \end{pmatrix}. \end{aligned}$$

It must be shown that $\operatorname{null} C = \{0\}$ and that $z^T B \overline{z} = 0$ and $C^T z = 0$ implies $z = 0$.

Firstly, $C\lambda = 0$ implies $P_{X'}\lambda = 0$ which implies $\lambda = 0$ since $\operatorname{null} P_{X'} = \{0\}$. Therefore $\operatorname{null} C = \{0\}$. Next, set $z^T = \left(v^T, \beta^T\right)$. Hence, since

$$C^T z = \begin{pmatrix} P_{X'}^T & O \end{pmatrix} \begin{pmatrix} v \\ \beta \end{pmatrix} = P_{X'}^T v,$$

we conclude that $C^T z = 0$ implies $P_{X'}^T v = 0$. Now assume that $C^T z = 0$. Then

$$\begin{aligned} z^T B \overline{z} = z^T \left((2\pi)^{\frac{d}{2}} N\rho G_{X';0} + \overline{G_P^T} G_P \right) \overline{z} &= (2\pi)^{\frac{d}{2}} N\rho\, v^T G_{X',X'} \overline{v}^T + z^T \overline{G_P^T} G_P \overline{z} \\ &= (2\pi)^{\frac{d}{2}} N\rho\, v^T G_{X',X'} \overline{v} + \left| G_P \overline{z} \right|_{\mathbb{C}^N}^2. \end{aligned}$$

By part 4 of Theorem 166, $G_{X',X'}$ is conditionally positive definite on $\operatorname{null} P_{X'}^T$ i.e. $P_{X'}^T v = 0$ and $v \neq 0$ implies $v^T G_{X',X'} \overline{v} > 0$. Therefore, $z^T B \overline{z} = 0$ implies $v = 0$.

In addition, we have $G_P\overline{z} = 0$ i.e. $0 = G_{X,X'}\overline{v} + P_X\overline{\beta} = P_X\overline{\beta}$, so that $\beta = 0$ since null $P_X = \{0\}$ by part 1 Theorem 104. We conclude that $z = 0$.

Part 2 We observe that $G_{X',X'}$ has size $N' \times N'$, $G_{X,X'}$ has size $N \times N'$,
P_X has size $N \times M$, and $P_{X'}$ has size $N' \times M$. From the block sizes it is clear that Ψ is square with $N' + 2M$ rows. Hence the size of Ψ is independent of the number of (scattered) data points N. ■

Our next result shows the Approximate smoother algorithm is *scalable* i.e. the time of execution of construction and solution is linearly dependent on the number of data points. This is in contrast with the Exact smoother which is not scalable but which has quadratic dependency on the number of data points.

Corollary 222 *The Approximate smoother algorithm is scalable.*

Proof. The Hermitian matrix equation 6.20 is

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + \overline{G^T_{X,X'}}G_{X,X'} & \overline{G^T_{X,X'}}P_X & P_{X'} \\ P^T_X G_{X',X} & P^T_X P_X & O_M \\ P^T_{X'} & O_M & O_M \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} = \begin{pmatrix} \overline{G^T_{X,X'}} \\ P^T_X \\ O_{M,N} \end{pmatrix} y,$$

where $G_{X',X'}$ is $N' \times N'$, $G_{X,X'}$ is $N \times N'$, $P_{X'}$ is $N' \times M$ and P_X is $N \times M$.

Suppose the *evaluation cost* for $G(x)$ is m_G multiplications and that $N \gg N'$, $N \gg M$ and $N \gg m_G$. The *construction costs* for component matrices and matrix multiplications are:

Construction costs for Approximate smoother matrix					
$G_{X',X'}$	$G_{X,X'}$	$\overline{G^T_{X,X'}}G_{X,X'}$	$P^T_X G_{X',X}$	$\Psi\begin{pmatrix}\alpha\\ \beta\\ \lambda\end{pmatrix}$	$\begin{pmatrix}\overline{G^T_{X,X'}}\\ P^T_X\\ O_{M,N}\end{pmatrix} y$
$(N')^2 m_G$	$N'Nm_G$	$(N')^2 N$	$N'NM$	$(N'+2M)^2$	$N(N'+2M)$

From the table the dominant *construction costs* are $N'N(N' + m_G + M)$. The *solution cost* of an $N' \times N'$ matrix equation is $\frac{1}{3}(N')^3$ multiplications for a *dense* matrix. Thus the total cost is

$$\begin{aligned} N'N(N' + m_G + M) + \frac{1}{3}(N')^3 &= \left((m_G + M)N' + (N')^2\right)N + \frac{1}{3}(N')^3 \\ &\gg \left((m_G + M)N' + (N')^2\right)N, \end{aligned}$$

which is linearly dependent on the number of data points. However, by the use of a basis function with support containing only several points in X' e.g. the extended natural spline basis functions of Lemma 53, the construction and solution costs can be reduced significantly. However we still have linear dependency on N. ■

6.5 Solving the Approximate smoothing problem using matrix techniques

In the last section, by using Hilbert space orthogonal projection techniques, the Approximate smoother matrix equation was derived for complex data and all conjugate-even basis functions. In this section we present an alternative derivation which uses matrix techniques. A key step in the Hilbert space approach was to choose a congugate-even basis function. This was sufficiently general and enabled us to prove that the Approximate smoother matrix was regular. This derivation is less general and imposes the **extra conditions** that the **basis function and data are real valued**. This is to simplify the use of the method of Lagrange multipliers and we note that the Approximate smoother matrix will now be symmetric and the basis function will be even.

The Approximate smoothing problem is $\min\limits_{f \in W_{G,X'}} J_e[f]$ where $J_e[f]$ is the Exact smoother functional. The first step will be to calculate $J_e[f]$ for $f \in W_{G,X'}$. The result is that $J_e[f]$ is a constrained quadratic form 6.21 in terms of the coefficients of the data-translated basis functions and the polynomial basis functions. Then it is shown that if the data is real and the basis function is real valued then the solution is unique and real valued. Finally, Lagrange multipliers are used to minimize the constrained quadratic form and derive the matrix equation as well as the identities derived using the Hilbert space method.

6.5.1 The Exact smoother functional for functions in $W_{G,X'}$

In this subsection we will calculate the Exact smoothing functional $J_e[f]$ for $f \in W_{G,X'}$.

Theorem 223 *Suppose the basis function G is a* ***real valued even function*** *and $f \in W_{G,X'}$, where $X' = \{x'_i\}_{i=1}^{N'}$ is unisolvent. Then if*

$$f = \sum_{i=1}^{N'} \alpha_i G(\cdot - x'_i) + \sum_{j=1}^{M} \beta_j p_j \in W_{G,X'}, \quad \alpha_i, \beta_j \in \mathbb{C},$$

we have

$$J_e[f] = \left(\alpha^T \; \beta^T\right)\left((2\pi)^{\frac{d}{2}} \rho G_{X';0} + \frac{1}{N} G_P^T G_P\right)\begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} - \frac{1}{N} y^T G_P \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \frac{1}{N} y^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} + \\ + \frac{1}{N}|y|^2. \tag{6.21}$$

where $J_e[\cdot]$ is the Exact smoother functional 5.6 and

$$G_{X';0} = \begin{pmatrix} G_{X',X'} & O \\ O & O_M \end{pmatrix}, \quad G_P = \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix}.$$

Proof. The Exact smoothing functional is

$$J_e[f] = \rho |f|_{w,\theta}^2 + \frac{1}{N} \sum_{i=1}^{N} \left| f\left(x^{(i)}\right) - y^{(i)} \right|^2,$$

and from part 1 of Theorem 166

$$\rho |f|_{w,\theta}^2 = \rho_\pi \alpha^T G_{X',X'} \alpha = \rho_\pi \left(\alpha^T \; \beta^T\right) G_{X';0} \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix}. \tag{6.22}$$

The next step is to approximate the second 'least squares' term $\sum\limits_{i=1}^{N} \left| f\left(x^{(i)}\right) - y^{(i)} \right|^2$. Let $f_X = \left(f\left(x^{(i)}\right)\right)_{i=1}^{N}$. Then

$$\sum_{i=1}^{N} \left| f(x^{(i)}) - y^{(i)} \right|^2 = (f_X - y)^T \left(\overline{f_X} - y\right) = f_X^T \overline{f_X} - y^T f_X - y^T \overline{f_X} + |y|^2. \tag{6.23}$$

But in matrix form

$$f_X = G_{X,X'} \alpha + P_X \beta,$$

where $G_{X,X'} = \left(G\left(x^{(i)} - x'_j\right)\right)$ and $P_X = \left(p_j\left(x^{(i)}\right)\right)$, or in block form

$$f_X = \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = G_P \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

where $G_P = \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix}$. Thus

$$f_X^T \overline{f_X} = \left(\alpha^T \; \beta^T\right) G_P^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix},$$

and by 6.23

$$\sum_{i=1}^{N} \left| f\left(x^{(i)}\right) - y^{(i)} \right|^2 = \left(\alpha^T \; \beta^T\right) G_P^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} - y^T G_P \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - y^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} + |y|^2, \tag{6.24}$$

so that combining 6.22 and 6.24

$$
\begin{aligned}
J_e[f] &= \rho |f|_{w,\theta}^2 + \frac{1}{N}\sum_{i=1}^{N}\left|f\left(x^{(i)}\right) - y^{(i)}\right|^2 \\
&= \rho_\pi \left(\alpha^T\, \beta^T\right) G_{X';0} \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} + \frac{1}{N}\left(\alpha^T\, \beta^T\right) G_P^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} - \\
&\qquad - \frac{1}{N} y^T G_P \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \frac{1}{N} y^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} + \frac{1}{N}|y|^2 \\
&= \left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} - \\
&\qquad - \frac{1}{N} y^T G_P \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \frac{1}{N} y^T G_P \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} + \frac{1}{N}|y|^2 .
\end{aligned}
$$

■

6.5.2 *Proof that the smoother is unique*

We will now supply another proof, this time based on matrices, that there *exists* a unique solution in $W_{G,X'}$ for the Approximate smoother problem. After that we will obtain a matrix equation for this solution.

Theorem 224 *The Approximate smoothing problem has a unique solution in $W_{G,X'}$.*

Proof. The Approximate smoother problem involves minimizing the quadratic form 6.21 constrained by $P_{X'}^T \alpha = 0$. But from part 2 of Theorem 105 the null space of $P_{X'}^T$ has the form $\alpha = \begin{pmatrix} -L_{X_2'}^T \\ I_{N'-M} \end{pmatrix} \alpha''$ where $\alpha'' = (\alpha_i)_{i=M+1}^{N'} \in \mathbb{R}^{(N'-M)}$ and $X_2' = \{x_i'\}_{i=M+1}^{N'}$. Set $A = \begin{pmatrix} -L_{X_2'}^T \\ I_{N'-M} \end{pmatrix}$ so that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} A & O \\ O & I_{N'} \end{pmatrix} \begin{pmatrix} \alpha'' \\ \beta \end{pmatrix}$ and the first term on the right of 6.21 becomes

$$
\begin{aligned}
&\left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} \\
&= \left(\alpha''^T\, \beta^T\right) \begin{pmatrix} A^T & O \\ O & I_{N'} \end{pmatrix} \left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} A & O \\ O & I_{N'} \end{pmatrix} \begin{pmatrix} \overline{\alpha''} \\ \overline{\beta} \end{pmatrix} .
\end{aligned} \tag{6.25}
$$

Thus the Approximate smoother problem is equivalent to minimizing 6.25 for all $\alpha'' \in \mathbb{R}^{(N'-M)}$ and $\beta \in \mathbb{R}^{N'}$. But if the matrix

$$
\begin{pmatrix} A^T & O \\ O & I_{N'} \end{pmatrix} \left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} A & O \\ O & I_{N'} \end{pmatrix}, \tag{6.26}
$$

is positive definite then the quadratic form 6.25 has a unique stationary point and this is a minimum point. Indeed

$$
\begin{aligned}
&\left(\alpha''^T\, \beta^T\right) \begin{pmatrix} A^T & O \\ O & I_{N'} \end{pmatrix} \left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} A & O \\ O & I_{N'} \end{pmatrix} \begin{pmatrix} \overline{\alpha''} \\ \overline{\beta} \end{pmatrix} \\
&\qquad = \left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N} G_P^T G_P\right) \begin{pmatrix} \overline{\alpha} \\ \overline{\beta} \end{pmatrix} \\
&\qquad = \rho_\pi \alpha^T G_{X',X'} \overline{\alpha} + \frac{1}{N}\left|\begin{pmatrix} G_{X,X'} & P_X \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right|_{\mathbb{C}^N}^2 \\
&\qquad = \rho_\pi \alpha^T G_{X',X'} \overline{\alpha} + \frac{1}{N}\left|G_{X,X'}\alpha + P_X \beta\right|_{\mathbb{C}^N}^2 ,
\end{aligned} \tag{6.27}
$$

where $P_{X'}^T \alpha = 0$. So the matrix 6.26 is positive definite if the matrix 6.27 is conditionally positive definite on null $P_{X'}^T$. So suppose

$$
\rho_\pi \alpha^T G_{X',X'} \overline{\alpha} + \frac{1}{N}\left|G_{X,X'}\alpha + P_X \beta\right|_{\mathbb{C}^N}^2 = 0, \quad P_{X'}^T \alpha = 0. \tag{6.28}
$$

Now by part 2 of Theorem 166, $G_{X',X'}$ is conditionally positive definite on $P_{X'}^T\alpha = 0$ i.e. when $P_{X'}^T\alpha = 0$ we have $\alpha^T G_{X',X'}\overline{\alpha} \geq 0$, and $P_{X'}^T\alpha = 0$ and $\alpha^T G_{X',X'}\overline{\alpha} = 0$ implies $\alpha = 0$. Thus $\alpha^T G_{X',X'}\overline{\alpha} = 0$ and $G_{X,X'}\alpha + P_X\beta = 0$ i.e. $\alpha = 0$ and $P_X\beta = 0$. But from part 1 of Theorem 104 $\operatorname{null} P_X = \{0\}$, and this implies $\beta = 0$. ■

The next result will simplify the use of Lagrange multipliers to solve the Approximate smoothing problem.

Theorem 225 *Suppose the basis function G is real-valued with order $\theta \geq 1$. Then the solution to the Approximate smoothing problem with real valued data $[X, y]$ induced by the points $X' = \{x_i'\}_{i=1}^{N'}$ lies in the subspace of $W_{G,X'}$ defined by the real scalars.*

Proof. The Approximate smoothing problem is $\min\limits_{f\in W_{G,X'}} J_e[f]$ and by the previous theorem this problem has a unique solution. Now let $f = \sum\limits_{i=1}^{N'} \alpha_i G(\cdot - x_i') + \sum\limits_{j=1}^{M} \beta_j p_j \in W_{G,X'}$, where $\alpha_i, \beta_j \in \mathbb{C}$. Then by Theorem 223

$$J_e[f] = \left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N}G_P^T G_P\right)\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} + \frac{1}{N}|y|^2,$$

and so

$$\begin{aligned}
J_e\left[\overline{f}\right] &= \left(\overline{\alpha}^T\, \overline{\beta}^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N}G_P^T G_P\right)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \\
&\quad + \frac{1}{N}|y|^2 \\
&= \left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N}G_P^T G_P\right)^T\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} + \\
&\quad + \frac{1}{N}|y|^2 \\
&= \left(\alpha^T\, \beta^T\right)\left(\rho_\pi G_{X';0} + \frac{1}{N}G_P^T G_P^T\right)\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \frac{1}{N}y^T G_P\begin{pmatrix}\overline{\alpha}\\ \overline{\beta}\end{pmatrix} + \\
&\quad + \frac{1}{N}|y|^2 \\
&= J_e[f].
\end{aligned}$$

■

Thus, if s is the Approximate smoother then $J_e[s] = J_e[\overline{s}]$ and $\overline{s}$ is also a solution to the Approximate smoothing problem. But the solution is unique so $s = \overline{s}$ and s is real-valued. Hence if $s = \sum\limits_{i=1}^{N'} \alpha_i G(\cdot - x_i') + \sum\limits_{j=1}^{M} \beta_j p_j$ then $s = \overline{s} = \sum\limits_{i=1}^{N'} \overline{\alpha_i} G(\cdot - x_i') + \sum\limits_{j=1}^{M} \overline{\beta_j} p_j$ and so $\sum\limits_{i=1}^{N'} (\alpha_i - \overline{\alpha_i}) G(\cdot - x_i') + \sum\limits_{j=1}^{M} \left(\beta_j - \overline{\beta_j}\right) p_j = 0$. The unique representation result of part 5 Theorem 166 now implies that for all i and j, $\alpha_i = \overline{\alpha_i}$ and $\beta_j = \overline{\beta_j}$ i.e. α_i and β_j are real. This means that $\sum\limits_{j=1}^{M} \beta_j p_j \in P_\theta$ and we are done.

6.5.3 The Approximate smoother matrix equation - a Lagrange multiplier derivation

Recall that the Approximate smoothing problem is $\min\limits_{f\in W_{G,X'}} J_e[f]$ and so this involves minimizing the quadratic form 6.21 constrained by $P_{X'}^T\alpha = 0$. In Theorem 224 it was proved that a solution to the Approximate smoothing problem exists and is unique. The proof also demonstrated that the constrained problem only has one stationary point which means the technique of Lagrange multipliers will always yield the Approximate smoother. For convenience denote the quadratic form 6.21 by $Q_a(\alpha, \beta)$ and denote the Lagrange multiplier by the vector λ with length N'. Since we have assumed the basis function is real-valued Theorem 225 tells us that the Approximate smoother can be expressed in terms of real α and β.

In preparation we expand the matrix equation 6.21 for $Q_a(\alpha,\beta)$ to include λ and for algebraic simplicity we will minimize NQ_a instead of Q_a. Thus for real vectors α and β, and $\rho_\pi = (2\pi)^{\frac{d}{2}}\rho$

$$NQ_a(\alpha,\beta) = \begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & O \\ O & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} -$$

$$- 2y^T \begin{pmatrix} G_P & O_{N,N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} + |y|^2, \quad \alpha \in \mathbb{R}^{N'}, \beta \in \mathbb{R}^M.$$

Let the Lagrangian be denoted by $L(\alpha,\beta,\lambda)$ so that

$$\begin{aligned}
L(\alpha,\beta,\lambda) &= NQ_a(\alpha,\beta) + 2\lambda^T P_{X'}^T \beta \\
&= NQ_a(\alpha,\beta) + \begin{pmatrix}\alpha^T \ \beta^T \ \lambda^T\end{pmatrix} \begin{pmatrix} O & O & P_{X'} \\ O & O & O \\ P_{X'}^T & O & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} \\
&= \begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & O \\ O & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} - \\
&\quad - 2y^T \begin{pmatrix} G_P & O_{N,N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} + |y|^2 + \\
&\quad + \begin{pmatrix}\alpha^T \ \beta^T \ \lambda^T\end{pmatrix} \begin{pmatrix} O_{N'} & O & P_{X'} \\ O & O_M & O \\ P_{X'}^T & O & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} \\
&= \begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} - \\
&\quad - 2y^T \begin{pmatrix} G_P^T & O_{N,N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} + |y|^2
\end{aligned} \tag{6.29}$$

A necessary condition for a minimum is that the derivative is zero i.e. $DL(\alpha,\beta,\lambda)=0$. Differentiating equation 6.29 gives

$$\begin{aligned}
DL &= \begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix} + \\
&\quad + \begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0}^T + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix}^T - 2y^T \begin{pmatrix} G_P & O \end{pmatrix} \\
&= 2\begin{pmatrix}\alpha^T \ \beta^T \vdots \lambda^T\end{pmatrix} \begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix} - 2y^T \begin{pmatrix} G_P & O \end{pmatrix}
\end{aligned}$$

At the minimum, $(DL)^T = 0$

$$\begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \cdots \\ \lambda \end{pmatrix} = \begin{pmatrix} G_P^T \\ O \end{pmatrix} y, \tag{6.30}$$

and since $G_P = \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix}$,

$$\begin{aligned}
N\rho_\pi G_{X';0} + G_P^T G_P &= N\rho_\pi G_{X';0} + \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix}^T \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix} \\
&= N\rho_\pi G_{X';0} + \begin{pmatrix} G_{X,X'}^T \\ P_X^T \end{pmatrix} \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix} \\
&= N\rho_\pi G_{X';0} + \begin{pmatrix} G_{X,X'}^T G_{X,X'} & G_{X,X'}^T P_X \\ P_X^T G_{X,X'} & P_X^T P_X \end{pmatrix} \\
&= \begin{pmatrix} N\rho_\pi G_{X',X'} & O \\ O & O \end{pmatrix} + \begin{pmatrix} G_{X,X'}^T G_{X,X'} & G_{X,X'}^T P_X \\ P_X^T G_{X,X'} & P_X^T P_X \end{pmatrix} \\
&= \begin{pmatrix} N\rho_\pi G_{X',X'} + G_{X,X'}^T G_{X,X'} & G_{X,X'}^T P_X \\ P_X^T G_{X,X'} & P_X^T P_X \end{pmatrix}
\end{aligned}$$

Thus

$$\begin{pmatrix} N\rho_\pi G_{X';0} + G_P^T G_P & \begin{matrix} P_{X'} \\ O \end{matrix} \\ \begin{matrix} P_{X'}^T & O \end{matrix} & O_{N'} \end{pmatrix} = \begin{pmatrix} N\rho_\pi G_{X',X'} + G_{X,X'}^T G_{X,X'} & G_{X,X'}^T P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix}, \tag{6.31}$$

and noting that $\rho_\pi = (2\pi)^{\frac{d}{2}} \rho$, equation 6.30 becomes

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X,X'}^T G_{X,X'} & G_{X,X'}^T P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} = \begin{pmatrix} G_{X,X'}^T \\ P_X^T \\ O \end{pmatrix} y,$$

which matches the Approximate matrix equation 6.20 which was derived using Hilbert space techniques and assumed a complex valued basis function.

6.6 Convergence of the Approximate smoother to the Exact smoother

In this section we will study the uniform pointwise convergence on a bounded region of the Approximate smoother s_a to the Exact smoother s_e and to the data function f_d.

In the first subsection we study the uniform convergence of the Approximate smoother to the Exact smoother. Here we make no use of Lagrange interpolation theory and obtain no order of convergence results.

In the second subsection we again consider the convergence of the Approximate smoother to the Exact smoother but here we use Lagrange interpolation theory and order of convergence results are derived.

The third subsection deals with the convergence of the Approximate smoother to the data function using the simple inequality

$$|f_d(x) - s_a(x)| \le |f_d(x) - s_e(x)| + |s_e(x) - s_a(x)|,$$

which combines the results of the previous subsection with those derived for the convergence of the Exact smoother to the data function in Section 5.7.

6.6.1 *Convergence results not involving order*

The convergence results of this subsection do not involve data functions and the dependent data is just given as an arbitrary vector. Also, no data densities are estimated and no orders of convergence are calculated, as will be done in Subsection 6.6.2. We simply establish uniform pointwise convergence on a bounded data region. To start with we will introduce a concept of convergence for finite subsets of $\mathbb{R}^d$.

Definition 226 ***Convergence of finite sets*** *A sequence of finite sets* $X_n = \left\{x_n^{(i)}\right\}_{i=1}^{N_n}$ *is said to converge to the finite set* $X = \left\{x^{(i)}\right\}_{i=1}^{N}$*, denoted* $X_n \to X$*, if there exists an integer* K *such that when* $n \ge K$*,* $N_n = N$ *and for each* i*,* $\left|x_n^{(i)} - x^{(i)}\right| \to 0$ *as* $n \to \infty$*.*

When $n \ge K$*,* X_n *and* X *can be regarded as members of* $\mathbb{R}^{Nd}$ *and convergence as convergence in* $\mathbb{R}^{Nd}$ *under the Euclidean norm.*

Theorem 227 *Suppose G is a basis function and s_e is the Exact smoother generated by the data $[X,y]$. Suppose X'_n is a sequence of independent data sets which converges to X_0 in the sense of Definition 226 and that $s_a^{(n)}$ is the Approximate smoother generated by X'_n and $[X,y]$.*

Then the Approximate smoothers satisfy $J_e\left[s_a^{(n)}\right] \to J_e[s_e]$ as $n\to\infty$, where J_e is the functional used to define the Exact smoother.

Proof. We first note that the definition of the convergence of independent data sets allows us to assume that the X'_n have the same number of points as X.

Now suppose X' is an arbitrary independent data set with the same number of points as X_0 and let $s_a = s_a(X')$ be the corresponding Approximate smoother. If it can be shown that as a function of X', $J_e[s_a(X')]$ is continuous everywhere this theorem holds since $J_e[s_e] = J_e[s_a(X)]$. In fact, by Theorem 218

$$s_a(X')(x) = \sum_{i=1}^{N} \alpha_i G(x - x'_i) + \sum_{j=1}^{M} \beta_j p_j(x),$$

where $\alpha = (\alpha_i)$ and $\beta = (\beta_j)$ satisfy the matrix equations

$$\begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} = \Psi^{-1} \begin{pmatrix} G^T_{X,X'} \\ P^T_X \\ O \end{pmatrix} y.$$

and the Approximate smoothing matrix is given by

$$\Psi = \begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G^T_{X,X'} G_{X,X'} & G^T_{X,X'} P_X & P_{X'} \\ P^T_X G_{X,X'} & P^T_X P_X & O_M \\ P^T_{X'} & O_M & O_M \end{pmatrix}.$$

Also, from part 4 Corollary 214 and then 6.15 we have

$$\begin{aligned} J_e[s_a(X')] &= \frac{1}{N} y^T (y - (s_a(X'))_X) \\ &= \frac{1}{N} y^T \left(y - \begin{pmatrix} G_{X,X'} & P_X \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right) \\ &= \frac{1}{N} y^T \left(y - \begin{pmatrix} G_{X,X'} & P_X & O \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} \right) \\ &= \frac{1}{N} y^T \left(I_N - \begin{pmatrix} G_{X,X'} & P_X & O \end{pmatrix} \Psi^{-1} \begin{pmatrix} G^T_{X,X'} \\ P^T_X \\ O \end{pmatrix} \right) y. \end{aligned}$$

If we can show that $G_{X,X'}$ and $\Psi^{-1}_{X'}$ are continuous functions of X' in a neighborhood of X_0 then $J_e[s_a(X')]$ is a continuous function of X'. But $G_{X',X'}$ and $G_{X,X'}$ are clearly continuous for all X' so $\Psi_{X'}$ and $\det \Psi_{X'}$ are continuous for all X'. Further, since by Theorem 221, $\Psi_{X'}$ is positive definite and regular for all X', $\det \Psi_{X'} > 0$ for all X' and it is clear from Cramer's rule that $\Psi^{-1}_{X'}$ is continuous everywhere. Thus $J_e[s_a(X')]$ is continuous everywhere and the proof is complete. ■

The next corollary shows that for given data the Approximate smoother converges uniformly to the Exact smoother on $\mathbb{R}^d$ for the case that X' is a grid containing a data region and the grid size goes to zero.

Corollary 228

1. *Suppose s_e is the Exact smoother generated by the data $[X,y]$ and that we have a sequence of regular, rectangular grids with grid nodes X'_n, grid sizes h'_n and a fixed common grid region $\mathcal{R}$ - a closed rectangle - containing X.*

2. *Let $s_a(X'_n)$ denote the Approximate smoother generated by X'_n and data $[X,y]$.*

3. *Assume that $X \subset \Omega \subseteq \mathcal{R}$ for all n, where the data region Ω is a bounded, open, connected subset of $\mathbb{R}^d$ having the cone property.*

4. *Assume that the conditions of Theorem 232 hold.*

Then $|h_n'| \to 0$ *implies* $\|s_a(X_n') - s_e\|_{w,\theta} \to 0$ *and also that* $s_a(X_n')$ *converges uniformly pointwise to* s_e *on* $\overline{\Omega}$.

Proof. For each data point $x^{(k)} \in X$ there exists a sequence of distinct points $\left(z_n^{(k)}\right)_{n=1}^{\infty}$ such that $z_n^{(k)} \in X_n'$ and $z_n^{(k)} \to x^{(k)}$ in $\mathbb{R}^d$ as $n \to \infty$.

Set $Z_n = \left\{z_n^{(k)}\right\}_{k=1}^{N}$ and let $s(Z_n)$ be the Approximate smoother generated by Z_n. Then $Z_n \to X$ as independent data and by Theorem 227, $J_e[s_a(Z_n)] \to J_e[s_e]$.

But since $Z_n \subset X_n'$ we have $J_e[s_a(X_n')] \le J_e[s_a(Z_n)]$.

Also from the definition of s_e we have $J_e[s_e] \le J_e[s_a(X_n')]$.

Thus

$$J_e[s_e] \le J_e[s_a(X_n')] \le J_e[s_a(Z_n)],$$

and so $J_e[s_a(X_n')] \to J_e[s_e]$.

To establish the convergence of $s_a(X_n')$ we use part 2 of Corollary 172, namely that

$$J_e[s_e] + \rho\,|s_e - f|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^{N}\left|s_e\left(x^{(k)}\right) - f\left(x^{(k)}\right)\right|^2 = J_e[f],$$

for all $f \in X_w^\theta$. Clearly $f = s_a(X_n')$ implies that as $n \to \infty$, $|s_e - s_a(X_n')|_{w,\theta} \to 0$ and $s_a(X_n')\left(x^{(k)}\right) \to s_e\left(x^{(k)}\right)$ for each k, and so

$$\|s_e - s_a(X_n')\|_{w,\theta}^2 = |s_e - s_a(X_n')|_{w,\theta}^2 + \sum_{m=1}^{M}|s_e(a_m) - s_a(X_n')(a_m)|^2 \to 0,$$

where the minimally unisolvent set of points $A = \{a_m\}_{m=1}^{M} \subset X$ define the Light norm $\|\cdot\|_{w,\theta}$. Finally, if R_x is the Riesz representer of the functional $f \to f(x)$ defined by A then for $x \in \Omega$

$$\begin{aligned}
|s_e(x) - s_a(X_n')(x)| &= \left|(s_e - s_a(X_n'), R_x)_{w,\theta}\right| \\
&\le \|s_e - s_a(X_n')\|_{w,\theta}\,\|R_x\|_{w,\theta} \\
&= \|s_e - s_a(X_n')\|_{w,\theta}\sqrt{|r_x|_{w,\theta}^2 + \left|\widetilde{l}(x)\right|^2} \\
&\le \|s_e - s_a(X_n')\|_{w,\theta}\left(|r_x|_{w,\theta} + \left|\widetilde{l}(x)\right|_1\right) \\
&= \|s_e - s_a(X_n')\|_{w,\theta}\left(\sqrt{r_x(x)} + \sum_{j=1}^{M}|l_j(x)|\right).
\end{aligned}$$

Our assumptions now allow us to choose A such that we can estimate $\sqrt{r_x(x)}$ and $\sum\limits_{j=1}^{M}|l_j(x)|$ using 6.33 and 6.37, and so obtain

$$|s_e(x) - s_a(X_n')(x)| \le \sqrt{c_{G,\eta,\sigma}}\left(1 + K_{\Omega,\theta}'\right)(\operatorname{diam}\Omega)^{\eta+\delta_G}\|s_e - s_a(X_n')\|_{w,\theta}, \quad x \in \Omega,$$

where $c_{G,\eta,\sigma}$ and $K_{\Omega,\theta}'$ are independent of $x \in \Omega$. Finally, since s_e and s_a are continuous on $\mathbb{R}^d$ the estimate is actually valid on $\overline{\Omega}$. Hence $s_a(X_n') \to s_e$ uniformly, pointwise on $\overline{\Omega}$ as $n \to \infty$. ■

The next corollary replaces the rectangular grids by scattered sets of points and thus it can be applied to sparse grids, for example. This corollary shows that for given data the Approximate smoother converges to the Exact smoother on $\mathbb{R}^d$ if the scattered sets converge to the independent data in the sense of Definition 226.

Corollary 229

1. *Suppose* s_e *is the Exact smoother generated by the data* $[X, y]$ *and that* $X \subset \Omega$ *where the data region* Ω *is a bounded, open, connected subset of* $\mathbb{R}^d$ *having the cone property.*

2. *Suppose that* X'_n *is a sequence of finite sets and that there exist a sequence of data sets* X''_n *such that* $X''_n \subset X'_n$ *and* $X''_n \to X$ *in the sense of Definition 226.*

3. *Assume that the conditions of Theorem 232 hold.*

Then, if $s_a(X''_n)$ *is the Approximate smoother generated by* $[X, y]$ *and* X''_n*, it follows that* $\|s_a(X''_n) - s_e\|_{w,\theta} \to 0$ *and* $s_a(X''_n)$ *converges uniformly pointwise to* s_e *on* $\overline{\Omega}$*.*

Proof. A simple modification of the proof of the previous Corollary 228. ■

6.6.2 Convergence to the Exact smoother

In Subsection 5.7.1 we proved results concerning the pointwise convergence of the Exact smoother to its data function using results from Lagrange interpolation theory and the 'cavity' measure of independent data set density $h_X = \sup_{\omega \in \Omega} \operatorname{dist}(\omega; X)$. In this subsection we will employ the same techniques to derive conditions under which the Approximate smoother converges pointwise to the Exact smoother - conditions expressed in terms of the smoothing coefficient ρ and the densities of the data sets X and X'.

We will require the following lemma which is derived from equation 5.29 and Theorem 187 of Chapter 5.

Lemma 230 *Suppose* $X = \{x^{(i)}\}_{i=1}^N$ *is a* θ*-unisolvent set and let* $A = \{a_i\}_{i=1}^M$ *be a minimally* θ*-unisolvent subset. Construct* R_x*,* $\mathcal{P}$ *and* $\mathcal{Q} = I - \mathcal{P}$ *using* A*.*

Then for each $x \in \mathbb{R}^d$ *there exists a unique element* $R_{V,x} \in V$ *such that*

$$f(x) = (\mathcal{L}_X f, R_{V,x})_V, \quad f \in X_w^\theta,$$

and

$$\rho \|R_{V,x}\|_V^2 \le r_x(x) + N_X \rho \left|\widetilde{l}(x)\right|^2. \tag{6.32}$$

We will also need the following lemma which supplies some results from the theory of Lagrange interpolation. This lemma has been created from Lemma 3.2, Lemma 3.5 and the first two paragraphs of the proof of Theorem 3.6 of Light and Wayne [10]. The results of this lemma do not involve any reference to weight or basis functions or to functions in X_w^θ, but consider the properties of the region Ω which contains the independent data points X and the order of the unisolvency used for the interpolation. Thus we have separated the part of the proof that involves basis functions from the part that uses the detailed theory of Lagrange interpolation operators.

Lemma 231 *(A copy of Lemma 148) Assume first that:*

1. Ω *is a bounded, open, connected subset of* $\mathbb{R}^d$ *having the cone property.*

2. X *is a unisolvent subset of* Ω *of order* θ*.*

Suppose $\{l_j\}_{j=1}^M$ *is the cardinal basis of* P_θ *with respect to a minimal unisolvent set of* Ω*. Using Lagrange interpolation techniques, it can be shown there exists a constant* $K'_{\Omega,\theta} > 0$ *such that*

$$\sum_{j=1}^M |l_j(x)| \le K'_{\Omega,\theta}, \quad x \in \overline{\Omega}, \tag{6.33}$$

and all minimal unisolvent subsets of Ω*. Now define*

$$h_X = \sup_{x \in \Omega} \operatorname{dist}(x, X),$$

and fix $x \in X$*. Using Lagrange interpolation techniques it can be shown there are constants* $c_{\Omega,\theta}, h_{\Omega,\theta} > 0$ *such that when* $h_X < h_{\Omega,\theta}$ *there exists a minimal unisolvent set* $A \subset X$ *satisfying*

$$\operatorname{diam} A_x \le c_{\Omega,\theta} h_X, \tag{6.34}$$

where $A_x = A \cup \{x\}$*.*

The next result supplies conditions on the basis function which may yield a higher order estimate for $\sqrt{r_x(x)}$ than Theorem 150.

Theorem 232 *(Copy of Theorem 156) Suppose w is a weight function with properties W2 and W3 for order θ and κ. Set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Also suppose G is a basis function of order θ such that the distributions $\left\{D^\beta G\right\}_{|\beta|=2\eta+1}$ are L^1_{loc} functions such that for each fixed $b \neq 0$ the integrals*

$$\int_0^1 (1-t)^{2\eta} \left|\left(D^\beta G\right)(z+tb)\right| dt, \quad z, b \in \mathbb{R}^d, \ |\beta| = 2\eta+1, \tag{6.35}$$

have polynomial growth in z.
Further, there exist constants $r_G, c_{G,\eta} > 0$ and $\delta_G \geq 0$ such that

$$\left|b^\beta\right| \int_0^1 (1-t)^{2\eta} \left|\left(D^\beta G\right)(tb)\right| dt \leq \frac{c_{G,\eta}}{2\sigma} |b|^{2(\eta+\delta_G)}, \quad |b| \leq r_G, \ |\beta| = 2\eta+1. \tag{6.36}$$

Regarding unisolvency, assume $A = \{a_k\}_{k=1}^M$ is a minimal θ-unisolvent set and that $\{l_k\}_{k=1}^M$ is the corresponding unique cardinal basis for P_θ. Now construct $\mathcal{P}, \mathcal{Q}, R_x$ using A and $\{l_k\}_{k=1}^M$.
Now if $r_x = \mathcal{Q}R_x$ we have the estimate

$$\sqrt{r_x(x)} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + \sum_{k=1}^M |l_k(x)|\right) (\operatorname{diam} A_x)^{\eta+\delta_G}, \quad \operatorname{diam} A_x \leq r_G, \ x \in \Omega, \tag{6.37}$$

where $\sigma = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa+1 \rfloor\right\}$, $A_x = A \cup \{x\}$ and $c_{G,\eta,\sigma} = \frac{d^{\lceil\sigma\rceil}}{(2\pi)^{d/2}\lceil\sigma\rceil!} c_{G,\eta}$.

Remark 233 *In Section 4.11, η and δ_G were calculated for the thin-plate and shifted thin-plate splines by way of examples.*

We now derive our estimates for the pointwise convergence of the Approximate smoother to the Exact smoother.

Theorem 234 *We will need the following assumptions and notation:*

(a) Suppose w is a weight function with properties W2 and W3 for order θ and κ, and set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Assume G is a basis function of order θ such that there exist constants $c_G, r_G > 0$ and $\delta_G \geq 0$ satisfying 6.37. Set $\eta_G = \eta + \delta_G$.

(b) Denote by s_e the Exact smoother generated by the the smoothing parameter ρ, the independent data X and data function $f_d \in X_w^\theta$.

(c) Assume $X \subset \Omega$ where the data region Ω has the properties given in the Lagrange interpolation Lemma 231 and the constants $h_{\Omega,\theta}, c_{\Omega,\theta}, K'_{\Omega,\theta}$ are as given in Lemma 231. Regarding Theorem 232 set

$$c_{G,\eta,\sigma} = \frac{d^{\lceil\sigma\rceil}}{(2\pi)^{d/2} \lceil\sigma\rceil!} c_{G,\eta}. \tag{6.38}$$

(d) Denote by s_a the Approximate smoother generated by the unisolvent set of points $X' \subset \Omega$. Let $h_X = \sup\limits_{\omega\in\Omega} \operatorname{dist}(\omega, X)$, $h_{X'} = \sup\limits_{\omega\in\Omega} \operatorname{dist}(\omega, X')$ measure the density of the point sets X and X'.

Then when $h_X, h_{X'} \leq \min\{h_{\Omega,\theta}, r_G\}$ we have for $x \in \overline{\Omega}$:

1. If $f_d \in W_{G,X'}$

$$|s_e(x) - s_a(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_{G,\eta,\sigma}} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} + \sqrt{N_X \rho}\right).$$

2. If $f_d \in X_w^\theta$

$$\begin{aligned} |s_e(x) - s_a(x)| \quad &\leq \quad |f_d|_{w,\theta} \left(1 + \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_{G,\eta,\sigma}} \frac{\left(c_{\Omega,\theta} h_{X'}\right)^{\eta_G}}{\sqrt{\rho}}\right) \times \\ &\quad \times \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_{G,\eta,\sigma}} \left(c_{\Omega,\theta} h_X\right)^{\eta_G} + \sqrt{N_X \rho}\right). \end{aligned} \tag{6.39}$$

Two similar bounds on the error for large ρ are:

3. *If* $f_d \in X_w^\theta$

$$|s_e(x) - s_a(x)| \leq |f_d|_{w,\theta} \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right)^2 (\operatorname{diam}\Omega)^{\eta_G} \left(\sqrt{c_{G,\eta,\sigma}} \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X}\right).$$

4. *If* $f_d \in X_w^\theta$

$$|s_e(x) - s_a(x)| \leq \left(\max_{\omega \in \overline{\Omega}} |f_d(\omega)|\right) \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_{G,\eta,\sigma}} \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X}\right).$$

5. *If* $f_d \in P_\theta$ *then* $f_d = s_e = s_a$.

Proof. Fix $x \in \Omega$ and construct $r_x = \mathcal{Q}R_x$, s_e and s_a from a minimal θ-unisolvent set $A \subset X$. From Lemma 230

$$|s_e(x) - s_a(x)| = \left|\left(\mathcal{L}_X(s_e - s_a), R_{V,x}\right)_V\right| \leq \|\mathcal{L}_X(s_e - s_a)\|_V \|R_{V,x}\|_V.$$

But from part 5 Summary 210, $\|\mathcal{L}_X(s_e - f)\|_V \leq \|\mathcal{L}_X f - \varsigma\|_V$ for all $f \in W_{G,X'}$. Choose $f = s_a$ so that by part 1 Summary 210

$$\|\mathcal{L}_X(s_e - s_a)\|_V^2 \leq \|\mathcal{L}_X s_a - \varsigma\|_V^2 = J_e[s_a],$$

and hence

$$|s_e(x) - s_a(x)| \leq \sqrt{J_e[s_a]} \|R_{V,x}\|_V.$$

But by 6.32

$$\rho \|R_{V,x}\|_V^2 \leq r_x(x) + N_X \rho \left|\widetilde{l}(x)\right|^2,$$

so that

$$|s_e(x) - s_a(x)| \leq \sqrt{J_e[s_a]} \frac{1}{\sqrt{\rho}} \left(\sqrt{r_x(x)} + \sqrt{N_X \rho} \left|\widetilde{l}(x)\right|\right), \quad x \in \mathbb{R}^d.$$

By Lemma 231 there exists a constant $K'_{\Omega,\theta}$, independent of $x \in \Omega$, such that $\left|\widetilde{l}(x)\right| \leq \sum\limits_{k=1}^{M} |l_k(x)| \leq K'_{\Omega,\theta}$ so that

$$|s_e(x) - s_a(x)| \leq \sqrt{J_e[s_a]} \frac{1}{\sqrt{\rho}} \left(\sqrt{r_x(x)} + K'_{\Omega,\theta} \sqrt{N_X \rho}\right), \quad x \in \Omega,$$

and the estimate 6.37 for $\sqrt{r_x(x)}$ on Ω implies

$$\sqrt{r_x(x)} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + \sum_{k=1}^{M} |l_k(x)|\right) (\operatorname{diam} A_x)^{\eta_G} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right) (\operatorname{diam} A_x)^{\eta_G},$$

provided $\operatorname{diam} A_x \leq r_G$. But an assumption of this theorem is $h_X \leq \min\{h_{\Omega,\theta}, r_G\}$ so by Lemma 231, A can be chosen so that $\operatorname{diam} A_x \leq c_{\Omega,\theta} h_X$ so that

$$\sqrt{r_x(x)} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right) (c_{\Omega,\theta} h_X)^{\eta_G}, \tag{6.40}$$

and

$$\begin{aligned} |s_e(x) - s_a(x)| &\leq \frac{\sqrt{J_e[s_a]}}{\sqrt{\rho}} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G} + \sqrt{N_X \rho}\right) \\ &= \left(1 + K'_{\Omega,\theta}\right) \frac{\sqrt{J_e[s_a]}}{\sqrt{\rho}} \left(\sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G} + \sqrt{N_X \rho}\right). \end{aligned}$$

Thus when $f \in W_{G,X'}$ and $f \neq s_a$, the definition of s_a implies $J_e[s_a] < J_e[f]$ and

$$|s_e(x) - s_a(x)| \leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{J_e[f]} \left(\sqrt{c_{G,\eta,\sigma}} \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X}\right). \tag{6.41}$$

The estimates for $|s_e(x) - s_a(x)|$ stated in the theorem will be proved by substituting various $f \in W_{G,X'}$ into 6.41.

Part 1 We choose $f = f_d \in W_{G,X'}$. Then $J_e[f] = \rho\,|f_d|_{w,\theta}^2$ and 6.41 implies

$$|s_e(x) - s_a(x)| \leq |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_{G,\eta,\sigma}}\,(c_{\Omega,\theta}h_X)^{\eta_G} + \sqrt{N_X\rho}\right), \quad x \in \Omega,$$

the extension to $\overline{\Omega}$ being valid since both s_e and s_a are continuous on $\mathbb{R}^d$.

Part 2 Choose $f = \mathcal{I}'_{X'}f_d$ where $\mathcal{I}'_{X'}f_d$ is the minimal seminorm interpolant of f_d on the unisolvent set X'. Then by Corollary 157, when $h_{X'} < \min\{h_{\Omega,\theta}, r_G\}$ it follows that

$$|f_d(x) - \mathcal{I}'_{X'}f_d(x)| \leq |f_d - \mathcal{I}'_{X'}f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\sqrt{c_{G,\eta,\sigma}}\,(c_{\Omega,\theta}h_{X'})^{\eta_G}, \quad x \in \overline{\Omega},$$

and so, by using the properties: $|\mathcal{I}'_{X'}f|_{w,\theta} \leq |f|_{w,\theta}$ and $|(I - \mathcal{I}'_{X'})f|_{w,\theta} \leq |f|_{w,\theta}$, of part 2 Theorem 147

$$\begin{aligned}
J_e[f] &= J_e[\mathcal{I}'_{X'}f_d] \\
&= \rho\,|\mathcal{I}'_{X'}f_d|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^{N}\left|(\mathcal{I}'_{X'}f_d)\left(x^{(k)}\right) - f_d\left(x^{(k)}\right)\right|^2 \\
&\leq \rho\,|\mathcal{I}'_{X'}f_d|_{w,\theta}^2 + |f_d - \mathcal{I}'_{X'}f_d|_{w,\theta}^2\left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma}\,(c_{\Omega,\theta}h_{X'})^{2\eta_G} \\
&\leq \rho\,|f_d|_{w,\theta}^2 + |f_d|_{w,\theta}^2\left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma}\,(c_{\Omega,\theta}h_{X'})^{2\eta_G} \\
&= |f_d|_{w,\theta}^2\left(\rho + \left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma}\,(c_{\Omega,\theta}h_{X'})^{2\eta_G}\right),
\end{aligned}$$

i.e.

$$\sqrt{J_e[f]} \leq |f_d|_{w,\theta}\left(\sqrt{\rho} + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_{G,\eta,\sigma}}\,(c_{\Omega,\theta}h_{X'})^{\eta_G}\right).$$

Hence by 6.41

$$\begin{aligned}
&|s_e(x) - s_a(x)| \\
\leq\;& \left(1 + K'_{\Omega,\theta}\right)\sqrt{J_e[f]}\left(\sqrt{c_{G,\eta,\sigma}}\frac{(c_{\Omega,\theta}h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X}\right) \\
\leq\;& |f_d|_{w,\theta}\left(\sqrt{\rho} + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_{G,\eta,\sigma}}\,(c_{\Omega,\theta}h_{X'})^{\eta_G}\right)\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_{G,\eta,\sigma}}\frac{(c_{\Omega,\theta}h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X}\right) \\
=\;& |f_d|_{w,\theta}\left(1 + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_{G,\eta,\sigma}}\frac{(c_{\Omega,\theta}h_{X'})^{\eta_G}}{\sqrt{\rho}}\right)\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_{G,\eta,\sigma}}\,(c_{\Omega,\theta}h_X)^{\eta_G} + \sqrt{N_X\rho}\right),
\end{aligned}$$

Part 3 Suppose $\mathcal{P}'$, $\mathcal{Q}'$ are r'_x are defined using a minimal unisolvent subset $A' \subset X'$. Choose $f = \mathcal{P}'f_d$. Then

$$\begin{aligned}
J_e[f] = J_e[\mathcal{P}'f_d] &= \rho\,|\mathcal{P}'f_d|_{w,\theta}^2 + \frac{1}{N}\sum_{k=1}^{N}\left|(\mathcal{P}'f_d)\left(x^{(k)}\right) - f_d\left(x^{(k)}\right)\right|^2 \\
&= \frac{1}{N}\sum_{k=1}^{N}\left|(\mathcal{Q}'f_d)\left(x^{(k)}\right)\right|^2 \\
&= \frac{1}{N}\sum_{k=1}^{N}\left|\left\langle f_d, r'_{x^{(k)}}\right\rangle_{w,\theta}\right|^2 \\
&\leq \frac{1}{N}\sum_{k=1}^{N}|f_d|_{w,\theta}^2\,|r'_{x^{(k)}}|_{w,\theta}^2 \\
&= |f_d|_{w,\theta}^2\,\frac{1}{N}\sum_{k=1}^{N}|r'_{x^{(k)}}|_{w,\theta}^2 \\
&\leq |f_d|_{w,\theta}^2\max_{\omega\in\overline{\Omega}}|r'_\omega|_{w,\theta}^2 \\
&= |f_d|_{w,\theta}^2\max_{\omega\in\overline{\Omega}}r'_\omega(\omega).
\end{aligned}$$

From 6.37

$$\sqrt{r_x(x)} \leq \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right) (\operatorname{diam}\Omega)^{\eta_G}, \quad x \in \Omega,$$

and since this is still true with A replaced by A' we have

$$\sqrt{J_e[f]} \leq |f_d|_{w,\theta} \max_{\omega \in \overline{\Omega}} \sqrt{r'_\omega(\omega)} \leq |f_d|_{w,\theta} \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right) (\operatorname{diam}\Omega)^{\eta_G},$$

and substitution into 6.41

$$\begin{aligned} |s_e(x) - s_a(x)| &\leq \left(1 + K'_{\Omega,\theta}\right) \sqrt{J_e[f]} \left(\sqrt{c_{G,\eta,\sigma}} \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X} \right) \\ &\leq |f_d|_{w,\theta} \sqrt{c_{G,\eta,\sigma}} \left(1 + K'_{\Omega,\theta}\right)^2 (\operatorname{diam}\Omega)^{\eta_G} \left(\sqrt{c_{G,\eta,\sigma}} \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{\rho}} + \sqrt{N_X} \right), \end{aligned}$$

which proves this part for $x \in \Omega$. Continuity on $\overline{\Omega}$ is valid since s_e and s_a are continuous on $\mathbb{R}^d$.

Part 4 Choose $f = 0$ so that $J_e[f] = J_e[0] \leq \left(\max_{\omega \in \overline{\Omega}} |f_d(\omega)| \right)^2$ and substitution into 6.41 yields this part for $x \in \Omega$. Continuity on $\overline{\Omega}$ is valid since s_e and s_a are continuous on $\mathbb{R}^d$.

Part 5 If $f_d \in P_\theta$ then by part 4, $J_e[s_a] = 0$ and so $|s_a|_{w,\theta} = 0$ and $s_a\left(x^{(k)}\right) = f_d\left(x^{(k)}\right)$ for all $x^{(k)} \in X$. Thus $s_a \in P_\theta$ and so $s_a = f_d$. But by property 2 of Theorem 175, $\mathcal{S}^e_X f = f$ iff $f \in P_\theta$ and consequently $s_e = f_d$, proving this part. ■

Remark 235

1. *The error formula of part 1 above is the same as the general Exact smoother error estimate of Theorem 236 Chapter 5. However, here it only applies to the special finite-dimensional subspace of data functions $W_{G,X'}$. Thus a sequence of independent data points X and smoothing coefficients ρ can be chosen so that the corresponding sequence of Approximate smoothers converges uniformly pointwise to the sequence of Exact smoothers, independently of the chosen X'.*

2. *The right-most factor of the estimate derived in part 2 of Theorem 234 is the estimate for the Exact smoother given below in Theorem 236. If we choose ρ such that*
 $\sqrt{N_X \rho} = \sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G}$ *i.e.*

$$\sqrt{\rho} = \frac{\sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{N_X}}, \tag{6.42}$$

 and then require that $\left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma} \frac{(c_{\Omega,\theta} h_{X'})^{\eta_G}}{\sqrt{\rho}} = 1$ *we get*

$$(c_{\Omega,\theta} h_{X'})^{\eta_G} = \frac{\sqrt{\rho}}{\left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma}} = \frac{\frac{\sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G}}{\sqrt{N_X}}}{\left(1 + K'_{\Omega,\theta}\right)^2 c_{G,\eta,\sigma}} = \frac{(c_{\Omega,\theta} h_X)^{\eta_G}}{\left(1 + K'_{\Omega,\theta}\right)^2 \sqrt{c_{G,\eta,\sigma} N_X}}.$$

 So if in addition

$$h_{X'} \leq \frac{h_X}{\left(\left(1 + K'_{\Omega,\theta}\right)^2 \sqrt{c_{G,\eta,\sigma} N_X} \right)^{1/\eta_G}}, \tag{6.43}$$

 then

$$|s_a(x) - s_e(x)| \leq 4 |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \sqrt{c_{G,\eta,\sigma}} (c_{\Omega,\theta} h_X)^{\eta_G}, \quad x \in \overline{\Omega}. \tag{6.44}$$

 Hence if the sequences X_k and X'_k are such that $h_{X_k} \to 0$ and $h_{X'_k} \to 0$ and ρ_k is constrained by 6.42 then there is a subsequence X'_{I_k} of X'_k such that $h_{X'_{I_k}} \leq \frac{h_{X_k}}{\left(\left(1+K'_{\Omega,\theta}\right)^2 \sqrt{c_{G,\eta,\sigma} N_{X_k}} \right)^{1/\eta_G}}$ and the sequence of Approximate smoothers converges to the sequence of Exact smoothers in the sense implied by 6.44.

 If X' is a regular grid then $h_{X'}$ can be calculated exactly:

$$d^{\frac{d}{2}} \operatorname{vol}(grid) = N_{X'} (h_{X'})^d. \tag{6.45}$$

3. *The approach of part 2 can be augmented by assuming a relationship between N_X and h_X. This was done in Section 4.8 Williams [22] for the zero order Exact smoother and in Remark 194 for the positive order Exact smoother. Several* ***1-dimensional*** *numerical experiments were run to compare the convergence of the zero order Exact smoother with the predicted convergence. 1-dimensional* ***test data sets*** *were constructed using a uniform distribution on the interval $\Omega = [-1.5, 1.5]$. Each of 20 data files were exponentially sampled using a multiplier of approximately 1.2 and a maximum of 5000 points, and then $\log_{10} h_X$ was plotted against $\log_{10} N_X$ where $N_X = |X|$. It then seemed quite reasonable to use a least-squares linear fit and in this case we obtained the relation*

$$h_X \simeq 3.09 N^{-0.81}. \tag{6.46}$$

For ease of calculation let

$$h_X = h_1 (N_X)^{-a}, \quad h_1 = 3.09, \quad a = 0.81. \tag{6.47}$$

A barrier to the use of such a formula as 6.46 in higher dimensions is the difficulty of calculating h_X for such data sets. If a sequence of independent test data sets was generated by a uniform distribution in each dimension then the constants a and h_1 might be defined as the upper bound of the confidence interval of a statistical distribution. Also, noting the regular grid formula 6.45, we might hypothesize a relationship of the form

$$h_X = h_d (N_X)^{-a_d d},$$

for ***higher dimensions****.*

4. *A similar approach, following Chapter 4 and Chapter 5 for the Exact smoother, is to substitute for N_X and $N_{X'}$ in 6.39 and minimize the estimator for ρ.*

6.6.3 *Convergence to the data function*

In the previous subsection we studied the convergence of the Approximate smoother to the Exact smoother. In this subsection we will combine these results with the Exact smoother error and thus estimate the error of the Approximate smoother. The relevant result concerning the Exact smoother error is Theorem 191 and our next result combines this result with Theorem 232 above:

Theorem 236

1. *The notation and assumptions of the Lagrange interpolation lemma (Lemma 231) hold.*

2. *Let w be a weight function with properties W2 and W3 for order θ and parameter κ and set $\eta = \min\left\{\theta, \frac{1}{2}\lfloor 2\kappa \rfloor\right\}$. Assume G is a basis function of order θ such that there exist constants $c_G, r_G > 0$ and $\delta_G \geq 0$ such that the estimate of the form 6.37 holds i.e.*

$$\sqrt{r_x(x)} \leq \sqrt{c_G} \left(1 + \sum_{k=1}^{M} |l_k(x)|\right) (\operatorname{diam} A_x)^{\eta_G}, \quad \operatorname{diam} A_x \leq r_G, \ x \in \Omega,$$

where $\eta_G = \eta + \delta_G$.

3. *Denote by s_e the Exact smoother generated by the the smoothing parameter ρ, the unisolvent independent data $X \subset \Omega$ and data function $f_d \in X_w^\theta$.*

Then when $h_X \leq \min\{h_{\Omega,\theta}, r_G\}$,

$$|s_e(x) - f_d(x)| \leq |f_d|_{w,\theta} \left(1 + K'_{\Omega,\theta}\right) \left(\sqrt{c_G} (c_{\Omega,\theta} h_X)^{\eta_G} + \sqrt{N_X \rho}\right), \quad x \in \overline{\Omega}. \tag{6.48}$$

We now present our main result for the convergence of the Approximate smoother to the data function.

Theorem 237 *We will use the assumptions and notation of Theorem 236 and Lemma 231. Now denote by s_a the Approximate smoother of the data function f_d generated by the data set X and the points X'. Let $h_{X'} = \sup_{\omega\in\Omega} \operatorname{dist}(\omega, X')$ measure the density of the point set X'.*

Then when $h_X, h_{X'} \leq \min\{h_{\Omega,\theta}, r_G\}$, $x \in \overline{\Omega}$ and $f_d \in X_w^\theta$ the estimate

$$\begin{aligned} & |f_d(x) - s_a(x)| \\ \leq \; & |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right)\left(2 + \left(1 + K'_{\Omega,\theta}\right)^2 c_G \frac{\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}}{\sqrt{\rho}}\right), \end{aligned} \tag{6.49}$$

holds.

Proof. From 6.39

$$\begin{aligned} |s_e(x) - s_a(x)| \; \leq \; & |f_d|_{w,\theta}\left(1 + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\frac{\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}}{\sqrt{\rho}}\right) \times \\ & \times \left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right), \quad x \in \overline{\Omega}, \end{aligned}$$

and from 6.48

$$|f_d(x) - s_e(x)| \leq |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right), \quad x \in \overline{\Omega},$$

so that

$$\begin{aligned} & |f_d(x) - s_a(x)| \\ \leq \; & |f_d(x) - s_e(x)| + |s_e(x) - s_a(x)| \\ \leq \; & |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right) + \\ & + |f_d|_{w,\theta}\left(1 + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\frac{\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}}{\sqrt{\rho}}\right)\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right) \\ = \; & |f_d|_{w,\theta}\left(1 + K'_{\Omega,\theta}\right)\left(\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \sqrt{N_X\rho}\right)\left(2 + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\frac{\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}}{\sqrt{\rho}}\right). \end{aligned}$$

■

Remark 238

1. *Comparison of the Approximate smoother error estimate proved in the last theorem with that of the estimate of part 2 of Theorem 234 shows that the convergence analysis given in Remark 235 can be applied to the estimate 6.49.*

2. *The Approximate smoother error estimate 6.49 is not bounded in ρ near zero or near infinity. However, estimates that are bounded in ρ near infinity can be derived as follows: First estimate $|s_e(x) - s_a(x)|$ by combining the estimate 6.39 with those of either part 3 or part 4 of the same theorem. Then add an Exact smoother error estimate from Theorem 201 to obtain the Approximate smoother error bounded in the smoothing parameter.*

3. *We now show that there exist sequences X'_k, X_k and ρ_k such that $\left|f_d(x) - s_a^{(k)}(x)\right| \to 0$ uniformly on $\overline{\Omega}$:*

   The estimator 6.49 has the form $(A + Bx)(C + D/x)$ where $x = \sqrt{\rho}$, $A = \sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}$, $B = \sqrt{N_X}$, $C = 2$ and $D = \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}$.

   The estimator is minimized to $\left(\sqrt{AC} + \sqrt{BD}\right)^2$ when $\rho = AD/BC$ i.e.

$$\begin{aligned} |f_d(x) - s_a(x)| \; & \leq \; \left(\sqrt{AC} + \sqrt{BD}\right)^2 \\ & \leq \; AC + BD \\ & = \; 2\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G} + \left(1 + K'_{\Omega,\theta}\right)\sqrt{c_G}\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}\sqrt{N_X}, \end{aligned}$$

when

$$\begin{aligned}\rho &= \frac{\sqrt{c_G}\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}\left(1+K'_{\Omega,\theta}\right)\sqrt{c_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}}{2\sqrt{N_X}} \\ &= \frac{1}{2}\left(1+K'_{\Omega,\theta}\right)\sqrt{c_G}\frac{\left(c_{\Omega,\theta}h_{X'}\right)^{\eta_G}\left(c_{\Omega,\theta}h_X\right)^{\eta_G}}{\sqrt{N_X}}.\end{aligned}$$

Hence if $\left(h_{X'_k}\right)^{\eta_G}\sqrt{N_{X_k}} \to 0$ *then* $\left|f_d(x) - s_a^{(k)}(x)\right| \to 0$.

4. ***These error estimates are quite unsatisfying! Phil must do better.***

6.7 A numerical implementation of the Approximate smoother

6.7.1 The SmoothOperator software (freeware)

In this section we discuss a numerical implementation for the construction and solution of the zero order Approximate smoother matrix equation 6.10 i.e.

$$\begin{pmatrix} (2\pi)^{\frac{d}{2}} N\rho G_{X',X'} + G_{X',X}G_{X,X'} & G_{X',X}P_X & P_{X'} \\ P_X^T G_{X,X'} & P_X^T P_X & O_M \\ P_{X'}^T & O_M & O_M \end{pmatrix} \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix} = \begin{pmatrix} G_{X',X} \\ P_X^T \\ O_{M,N} \end{pmatrix} y,$$

I called this software *SmoothOperator*. In Corollary 222 it was shown that the construction and solution of this matrix equation is a scalable process and thus worthy of numerical implementation. This software also implements the zero order version of this equation which is derived in Williams [22] and the tutorials concentrate on the zero order basis functions.

The algorithm has been implemented in **Matlab 6.0** with a GUI interface but has only been tested on Windows. *SmoothOperator* can be obtained by emailing the author. However there is a **short user document** (4 pages) and the potential user can read this document first to decide whether they want the software. The top-level directory of the software contains the file *read_me.txt* which can also be downloaded separately. A **full user manual** (82pages) comes with the software. The main features of the full user manual are:

1 Tutorials and data experiments To learn about the system and the behavior of the algorithm, I have prepared three tutorials and five data experiments.

2 Context-sensitive help Each dialog box incorporates context-sensitive help which is invoked using the right mouse button. An F1 key facility could be implemented. The actual help text is contained in the text file **\Help\ContextHelpText.m** that can be easily edited.

3 Matlab diary facility When the system is started the Matlab diary facility is invoked. This means that most user information generated by *SmoothOperator* and written to the Matlab command line is also written to the text diary file. Each dialog box has a drop-down diary menu which allows the user to view the diary file using Notepad. There is also a facility for you to choose another editor/browser. The file can also be emptied, if, for example, it gets too big. It can also be disabled.

4 Tools for viewing data files Before generating a smoother you can view the contents of the data file.

For ASCII delimited text data files, use the (slow) **View records** facility to display records and the file header, and then with this information you can use the high speed **Study records** facility to check the records and then obtain detailed information about single fields e.g. a histogram, and multiple fields e.g. correlation coefficients and scatter plots.

For binary **test data files** only the **View records** facility is needed. The parameters which generated the file can also be viewed using the **Make or View data** option.

5 The output data The output from the experiments and tutorials mentioned in point 1 above consists of well-documented Matlab one and two-dimensional plots, command line output and diary output. There is currently no file output, but this can be implemented on request.

6 Reading delimited text files A MEX C file allows ASCII delimited text files to be read very quickly. You can specify:

6.1 that the file be read in chunks of records, and not in one go.
6.2 The ids of the fields to be read. This means that the file can contain non-numeric fields.
6.3 Fields can be checked to ensure they are numeric.

The tutorials which create smoothers allow the data, smoothed data, and related functions to be viewed using scatter plots and plots along lines and planes.

7 Please note that there is no explicit suite of functions - application programmer interface or API - supplied by *SmoothOperator* for immediate use by the user. After reflecting on how I would create such an API, I decided that I lacked the experience to produce it, and that, anyhow, there were too many possibilities to anticipate. I would expect that possible users of this software, designed to be applied to perhaps millions of records, would need to familiarize themselves thoroughly with how this system works. I urge them to contact me and discuss their application.

6.7.2 Algorithms

The following three algorithms are used in the *SmoothOperator* software package to calculate the Approximate smoother.

Algorithm 1 uses data generated internally according to specifications supplied by the user using the interface. The smoothing parameter can be either specified or calculated using an error grid. This algorithm is **scalable**.

Algorithm 2 uses a 'small' subset of actual data to get an idea of a suitable smoothing parameter to use for the full data set. This algorithm is **not scalable**.

Algorithm 3 uses all the actual data. The smoothing parameter can be either specified or calculated using an *error grid*. This algorithm is **scalable**.

We will now explain these algorithms in more detail.

Algorithm 1: using experimental data

1. Generate the experimental data $[X, y]$. The independent data X is generated by uniformly distributed random numbers on X. The dependent data y is generated by uniformly perturbing an analytic data function g_{dat}.

2. Choose a smoothing grid X' whose boundary contains X. Choose an error grid X'_{err} which will be used to estimate the optimal smoothing parameter ρ.

3. Read the data $[X, y]$ and construct the matrix equation. If the matrix $G_{X,X'}$ is dense, the matrix $G^T_{X,X'}G_{X,X'}$ is constructed using a Matlab mex file (a compiled C file). If $G_{X,X'}$ is sparse the usual matrix multiplication is used.

4. Given a value for the smoothing parameter ρ we can solve the matrix equation and evaluate the smoother.

   We want to estimate the value of ρ which minimizes the 'sum of squares' error between the smoother and the data function

   $$\delta_1(\rho) = \sum_{x'' \in X'_{err}} \left(\sigma_\rho(x'') - g_{dat}(x'')\right)^2, \quad \rho > 0.$$

   Empirical work indicates a standard shape for $\delta_1(\rho)$, namely decreasing from right to left, reaching a minimum and then increasing at a decreasing rate. To find the minimum we basically use the standard iterative algorithm of dividing and multiplying by a factor e.g. 10 and choosing the smallest value. The process is stopped when the percentage change of one or both of $\delta_1(\rho)$ and ρ are less than prescribed values.

Algorithm 2: using a 'test' subset of actual data

1. Perhaps based on results using Algorithm 1, choose an initial value for smoothing parameter ρ and choose a smoothing grid X'.

2. Step 2 of Algorithm 1. Denote the data by $[X, y]$ where $X = \left(x^{(i)}\right)$ and $y = (y_i)$.

3. This is the same as step 4 of Algorithm 2 except we now minimize

$$\delta_2(\rho) = \sum_{i=1}^{N} \left(\sigma_\rho\left(x^{(i)}\right) - y_i\right)^2, \quad \rho > 0, \tag{6.50}$$

because we do not have a data function.

Algorithm 3: using all the actual data

1. Choose a value for smoothing parameter ρ, based on experiments using Algorithm 2. Choose a smoothing grid X'.
2. Read all the data $[X, y]$ and construct the matrix equation. If the matrix $G_{X,X'}$ is dense, the matrix $G^T_{X,X'}G_{X,X'}$ is constructed using a Matlab mex file (a compiled C file). If $G_{X,X'}$ is sparse the usual multiplication is used.
3. Solve the matrix equation to obtain the basis function coefficients and evaluate the smoother at the desired points.

6.7.3 Features of the smoothing algorithm and its implementation

1 The Short user manual and the User manual contains a lot of detail regarding the *SmoothOperator* system and algorithms. So we will content ourselves here with just some key points.

2 Although the algorithm is scalable there can still be a problem with rapidly increasing memory usage as the grid size decreases and the dimension increases. The classical radial basis functions, such as the thin plate spline functions, have support everywhere. Hence the smoothing matrix is completely full. To significantly reduce this problem we do the following:
a) We use basis functions with bounded support.
b) We shrink the basis function support to the magnitude of the grid cells. This makes $G^T_{X,X'}G_{X,X'}$ a **very sparse** banded diagonal matrix, and $G_{X',X'}$ has a small number of non-zero diagonals. We say this basis function has *small support*.

3 Instead of using the memory devouring Matlab `repmat` function to calculate $G_{X,X'}$ and $G_{X',X'}$, we directly calculate the arguments of the Matlab function `sparse`. This function takes three arrays, namely the row ids, the column ids and the corresponding matrix elements, as well as the matrix dimensions, and converts them to the Matlab sparse internal representation. I must mention that although this is quick and space efficient, the algorithm is much more complicated than using `repmat`.
Matlab's sparse multiplication facility is then used to quickly and efficiently calculate $G^T_{X,X'}G_{X,X'}$.

4 A selection of basis functions is supplied.

5 Note that this software was written before I embarked on the error analysis contained in this document so *SmoothOperator* concentrates on the Approximate smoother and the user cannot study the error of the Exact smoother or the Approximate smoother or compare the Approximate smoother with the Exact smoother.

6.7.4 An application - predictive modelling of forest cover type

In this section we demonstrate how the method developed in the previous sections can be used in data mining for predictive modelling. This application smooths **binary-valued** data - I was unable to obtain a large 'continuous-valued' data set for distribution on the web. The source of our data is the web site file:

http://kdd.ics.uci.edu/databases/covertype/covertype.html,

in the UCI KDD Archive, Information and Computer Science, University of California, Urvine.
The data gives the forest cover type in 30×30 meter cells as a function of the following cartographic parameters:
Forest cover type is the dependent variable y and it takes on one of seven values:

Id	Independent variable	Description
1	ELEVATION	Altitude above sea level
2	ASPECT	Azimuth
3	SLOPE	Inclination
4	HORIZ_HYDRO	Horizontal distance to water
5	VERT_HYDRO	Vertical distance to water
6	HORIZ_ROAD	Horizontal distance to roadways
7	HILL_SHADE_9	Hill shade at 9am
8	HILL_SHADE_12	Hill shade at noon
9	HILL_SHADE_15	Hill shade at 3pm
10	HORIZ_FIRE	Horizontal distance to fire points

TABLE 6.1.

Forest cover type	Id
Spruce fir	1
Lodge-pole pine	2
Ponderosa pine	3
Cottonwood/Willow	4
Aspen	5
Douglas fir	6
Krummholtz	7

TABLE 6.2.

The data file

In this study we will use the data to train a model predicting on the **presence or absence of the Ponderosa pine forest cover** (id = 3), but the results will be similar for the other forest types. To this end we have created from the full web site file a file called `\UserData\forest_1_to_10_pondpin.dat` which contains the ten independent variables of Table 6.1 and then a binary dependent variable derived from the variable **Ponderosa pine** of Table 6.2. This variable is 1 if the cover is Ponderosa pine and zero otherwise.

Methodology

We recommend you first use the user interface of the *SmoothOperator* software to construct artificial data sets to understand the behavior of the smoother and run the experiments to get a feel for the influence of the parameters.

1 We chose a small subset of the forest-cover data and selected various smoothing parameters and smoothing grid sizes to study the performance of the smoother using plots and the value of the error.

Note that *two subsets of data* could be used here, often called the *training set* and the *test set*. The training set would be used to calculate the smoother for the initial value of the smoothing coefficient, and then the test set could be used to determine the smoothing coefficient which minimizes the least squares error.

2 Having chosen our parameters we run the smoother program on the full data set.

Appendix A

Basic notation, definitions and symbols

A.1 Basic function and distribution spaces

Definition 239 ***Basic function and distribution spaces***
All spaces below consist of ***complex-valued*** *functions:*

1. $P_0 = \{0\}$. *For* $n \geq 1$, P_n *denotes the polynomials of (total) order at most* n *i.e. degree at most* $n-1$. *These polynomials have the form* $\sum\limits_{|\alpha|<n} a_\alpha \xi^\alpha$, *where* $a_\alpha \in \mathbb{C}$ *and* $\xi \in \mathbb{R}^d$. *The space of all polynomials will be denoted by* P.

2. $C^{(0)}$ *is the space of continuous functions.*

   $C_B^{(0)}$ *is the space of bounded continuous functions.*

   $C_{BP}^{(0)}$ *is the space of continuous functions bounded by a polynomial.*

   $C^{(m)} = \left\{f \in C^{(0)} : D^\alpha f \in C^{(0)},\ \text{when } |\alpha| = m\right\}$.

   $C_B^{(m)} = \left\{f \in C_B^{(0)} : D^\alpha f \in C_B^{(0)},\ \text{when } |\alpha| \leq m\right\}$.

   $C_{BP}^{(m)} = \left\{f \in C_{BP}^{(0)} : D^\alpha f \in C_{BP}^{(0)},\ \text{when } |\alpha| \leq m\right\}$.

   $C^\infty = \bigcap\limits_{m\geq 0} C^{(m)}$; $C_B^\infty = \bigcap\limits_{m\geq 0} C_B^{(m)}$; $C_{BP}^\infty = \bigcap\limits_{m\geq 0} C_{BP}^{(m)}$.

3. L_{loc}^1 *is the space of measurable functions which are absolutely integrable on any compact set i.e. any closed, bounded set.*

   L^1 *is the Hilbert space of measurable functions* f *such that* $\int |f| < \infty$. *Norm is* $\|f\|_1 = \int |f|$ *and inner product is* $(f,g)_1$.

   L^2 *is the Hilbert space of measurable functions* f *such that* $\int |f|^2 < \infty$. *Norm is* $\|f\|_2 = \left(\int |f|^2\right)^{1/2}$ *and inner product is* $(f,g)_2$.

   L^∞ *is the space of a.e. bounded functions. Norm is* $\|f\|_\infty$.*and inner product is* $(f,g)_\infty$.

4. C_0^∞ *is the space of* C^∞ *functions that have compact support. These are the test functions for the space of distributions defined on* $\mathbb{R}^d$, *sometimes denoted by* $\mathcal{D}$.

   S *is the* C^∞ *space of rapidly decreasing functions. A function* $f \in C^\infty$ *is in* S *if given any multi-index* α *and integer* $n \geq 0$, *there exists a constant* $k_{\alpha,n}$ *such that,* $|D^\alpha f(x)| \leq k_{\alpha,n}(1+|x|)^{-n}$, $x \in \mathbb{R}^d$. *These are the test functions for the tempered distributions of Definition 243 below.*

   $\mathcal{D}'$ *is the space of distributions (generalized functions).*

   $\mathcal{E}'$ *is the space of distributions with compact (bounded) support.*

A.2 Vector notation

Definition 240 ***Vector notation***
Suppose v and w are real vectors:

- $\mathbf{0} = (0, 0, \ldots, 0)$ *and* $\mathbf{1} = (1, 1, \ldots, 1)$.
- *Suppose* $\tilde{}$ *is one of the binary operations* $<, \leq, =, >, \geq$.
  Then we write $v\tilde{}w$ *if* $v_i\tilde{}w_i$ *for all* i.
  For $x \in \mathbb{R}$ *we write* $v\tilde{}x$ *if* $v_i\tilde{}x$ *for all* i.
- $v^w = (v_1)^{w_1} (v_2)^{w_2} \ldots (v_d)^{w_d}$.
- *If* $s \in \mathbb{R}$ *then* $v^s = v^{s\mathbf{1}}$.

A.3 Topology

Definition 241 ***Topology on*** R^d

- *The Euclidean norm an inner product are denoted by* $|x|$ *and* (x, y).
- *Ball* $B(x; r) = \{y : |x - y| < r\}$.
- $\varepsilon-$*neighborhood of a set - for* $\varepsilon > 0$, *the* $\varepsilon-$*neighborhood of a set* S *is denoted* S_ε *and* $S_\varepsilon = \bigcup_{x \in S} B(x; \varepsilon)$.

A.4 Multi-indexes

Definition 242 ***Multi-indexes***

1. *Multi-indexes are vectors with non-negative integer components.*
   Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ *and* $\beta = (\beta_1, \beta_2, \ldots, \beta_d)$ *be multi-indexes.*
   Suppose $\tilde{}$ *is one of the binary operations* $<$, $\leq$, $=$, $>$, $\geq$.
   Write $\beta\tilde{}a$ *if* $\beta_i\tilde{}\alpha_i$ *for all* i.
   For $x \in \mathbb{R}$ *write* $\beta\tilde{}x$ *if* $\beta_i\tilde{}x$ *for all* i.
2. *Denote* $|\alpha| = \sum_{i=1}^{d} \alpha_i$. *Then* $D^\alpha f(x)$ *is the derivative of the function* f *of degree* α
   $$D^0 f(x) = f(x), \quad D^\alpha f(x) = \frac{D^{|\alpha|} f(x_1, x_2, \ldots, x_d)}{D_1^{\alpha_1} x_1 D_2^{\alpha_2} x_2 \ldots D_d^{\alpha_d} x_d}.$$
3. *We shall also use the notation*
   $$\begin{array}{ll} \textit{monomial} & x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_d^{\alpha_d}, \\ \textit{factorial} & \alpha! = \alpha_1!\alpha_2! \ldots \alpha_d! \textit{ and } 0! = 1, \\ \textit{binomial} & \dbinom{\alpha}{\beta} = \dfrac{\alpha!}{(\alpha - \beta)!\beta!}, \quad \textit{if } \beta \leq \alpha. \end{array}$$
4. *The inequality* $|x^\alpha| \leq |x|^{|\alpha|}$ *is used often.*
5. *Important identities are*
   $$(x, y)^k = \sum_{|\alpha| = k} \frac{k!}{\alpha!} x^\alpha y^\alpha, \quad |x|^{2k} = \sum_{|\alpha| = k} \frac{k!}{\alpha!} x^{2\alpha}.$$
6. *Leibniz's rule is*
   $$D^\alpha (uv) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta u D^{\alpha - \beta} v. \tag{A.1}$$

A.5 Tempered distributions

Definition 243 ***Tempered distributions*** S' *(also temperate or generalized functions of slow growth)*

1. *S is the space of rapidly decreasing functions. We endow S with the topology defined using the countable set of seminorms $p_{n,\alpha}$ given by*

$$p_{n,\alpha}(\psi) = \left\| (1 + |\cdot|)^n D^\alpha \psi \right\|_\infty, \quad n = 0, 1, 2, \ldots; \ \alpha \geq 0.$$

2. *S' is the space of tempered distributions or generalized functions of slow growth. It is the set of all continuous linear functionals on S under the seminorm topology of part 1 of this definition.*

3. *A linear functional on S is continuous if its absolute value is bounded by a finite positive linear combination of seminorms.*

4. *If $f \in S'$ and $\phi \in S$ then $[f, \phi] \in \mathbb{C}$ will represent the action of f on the test function ϕ.*

A.5.1 Properties

1. The continuous embeddings $S \subset C_{BP}^\infty \subset S' \subset \mathcal{D}'$ are dense.

2. If $f \in L_{loc}^1$ and, $\int |f(x)| (1 + |x|)^{-\lambda} dx$ or $\int_{|\cdot| \geq r} |f(x)| |x|^{-\lambda} dx$ exist for some $\lambda, r \geq 0$, then $f \in S'$ with action $[f, \phi] = \int f(x) \phi(x) dx$, $\phi \in S$. We call f a *regular tempered distribution (function)* e.g. Example 2.8.3(a) Vladimirov [21].

3. If $f \in S'$ and $a \in C_{BP}^\infty$ then $af \in S'$ e.g. Example 2.8.3(e) Vladimirov [21] but note that Vladimirov denotes C_{BP}^∞ by θ_M.

A.6 Fourier Transforms

Definition 244 ***Fourier and Inverse transforms on*** S ***and*** S'***.***

1. *If $f \in S$ then let the Fourier transform be*

$$\widehat{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-ix\xi} f(x) dx.$$

Where convenient the alternative notation $F[f] = \widehat{f}$ will be used.

The mapping $f \to \widehat{f}$ is continuous from $S \to S$.

The inverse Fourier transform is

$$\overset{\vee}{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{ix\xi} f(x) dx.$$

Where convenient the alternative notation $F^{-1}[f] = \overset{\vee}{f}$ will be used.

We have the important property that $\left(\widehat{f}\right)^\vee = f$.

2. *If $f \in S'$ and $\phi \in S$ then $\left[\widehat{f}, \phi\right] = \left[f, \overset{\vee}{\phi}\right]$ and $\left[\overset{\vee}{f}, \phi\right] = \left[f, \overset{\vee}{\phi}\right]$.*

The mappings $f \to \widehat{f}$ and $f \to \overset{\vee}{f}$ are continuous from $S' \to S'$.

We have the important property that $\left(\widehat{f}\right)^\vee = f$.

A.6.1 Fourier transform properties on S and S'

1. If b is a complex constant then $(f(x+b))^{\wedge} = e^{ib\xi}\widehat{f}$.

2. $\overset{\vee}{f}(\xi) = \widehat{f}(-\xi)$, $\overline{\widehat{f}}(\xi) = \overset{\vee}{\overline{f}}(\xi)$ and $\widehat{\widehat{f}}(\xi) = f(-\xi).\overline{\hat{f}}$

3. If b is a complex constant then $\left(e^{ibx}\right)^{\wedge} = (2\pi)^{d/2}\,\delta(\xi+b)$.

4. $D^{\alpha}\widehat{f} = (-i)^{|\alpha|}\,\widehat{x^{\alpha}f}$ and $D^{\alpha}\overset{\vee}{f} = i^{|\alpha|}\,(x^{\alpha}f)^{\vee}$.

5. $\widehat{x^{\alpha}f} = i^{|\alpha|}D^{\alpha}\widehat{f}$.

6. $\widehat{D^{\alpha}f} = i^{|\alpha|}\xi^{\alpha}\widehat{f} = (i\xi)^{\alpha}\,\widehat{f}$.

7. $\xi^{\alpha}\widehat{f} = (-i)^{|\alpha|}\,\widehat{D^{\alpha}f}$.

8. $\widehat{\delta} = (2\pi)^{-d/2}$ and $\widehat{1} = (2\pi)^{d/2}\,\delta$.

9. $\widehat{x^{\alpha}} = (2\pi)^{d/2}\,(iD)^{\alpha}\,\delta$ and $\widehat{D^{\alpha}\delta} = (2\pi)^{-d/2}\,(-i)^{|\alpha|}\,\xi^{\alpha}$.

10. If p is a polynomial then $\widehat{p} = (2\pi)^{d/2}\,p(iD)\delta$, and $\widehat{p(D)\,f} = p(-i\xi)\,\widehat{f}$.

A.6.2 Fourier transform properties on $\mathcal{E}'$

1. Suppose $f \in \mathcal{E}'$, $\eta \in C_0^{\infty}$ and $\eta = 1$ in a neighborhood of the support of f - denoted $\operatorname{supp} f$. Then f can be extended uniquely to S' by $[f,\phi] = [f,\eta\phi]$, where $\phi \in S$. e.g. Example 2.8.3(b) Vladimirov [21] but note that Vladimirov denotes C_{BP}^{∞} by θ_M.

2. If $f \in \mathcal{E}'$ then $f \in S'$ in the sense of part 1and $\widehat{f} \in C_{BP}^{\infty}$ e.g. Theorem 2.9.4 Vladimirov [21].

A.7 Convolutions

Definition 245 ***Convolution***

1. *If $f \in C_{BP}^{(0)}$ and $\phi \in S$ then $f * \phi \in C_{BP}^{\infty}$ where*

$$(f*\phi)(x) = (\phi*f)(x) = (2\pi)^{-\frac{d}{2}}\int f(x-y)\,\phi(y)\,dy = (2\pi)^{-\frac{d}{2}}\int f(y)\,\phi(x-y)\,dy.$$

2. *The previous definition can be extended to S' by the formulas*

$$f*\phi = \phi*f = \left(\widehat{\phi}\widehat{f}\right)^{\vee} = (2\pi)^{-\frac{d}{2}}\left[f_y, \phi(\cdot - y)\right], \quad f \in S',\ \phi \in S.$$

   Here $f*\phi \in C_{BP}^{\infty}$.

3. *Noting parts 1 and 3 of Appendix A.5.1, the following definition is consistent with parts 1 and 2:*

$$f*g = g*f = \left(\widehat{g}\widehat{f}\right)^{\vee}, \quad f \in S',\ g \in (C_{BP}^{\infty})^{\wedge}.$$

   Here $f*\phi \in S'$.

4. *Also, noting the results of Appendix A.6.2*

$$f*g = g*f = \left(\widehat{g}\widehat{f}\right)^{\vee}, \quad f \in S',\ g \in \mathcal{E}'.$$

   Here $f*g \in S'$

5. *If $f \in C_{BP}^{(0)}$ and $g \in C_0^{(0)}$ then $f*g$ is a regular tempered distribution and*

$$(f*g)(x) = (g*f)(x) = (2\pi)^{-\frac{d}{2}}\int f(x-y)\,g(y)\,dy = (2\pi)^{-\frac{d}{2}}\int f(y)\,g(x-y)\,dy.$$

   This can be proved using Theorem 2.7.5 of Vladimirov [21].

A.7.1 Convolution properties

For the convolutions defined in parts 2 and 4 above:

1. $f * g = g * f$.
2. $D^{\alpha}(f * g) = f * D^{\alpha} g = D^{\alpha} f * g$ for all α.
3. $(f * g)^{\wedge} = \widehat{g}\widehat{f}, \quad$ where $\widehat{f} \in S'$.
4. The mapping $f \to f * g$ is continuous from S' to S'.

A.8 Taylor series expansion

Suppose $u : \mathbb{R}^d \to \mathbb{C}$ and $u \in C^{(n)}\left(\mathbb{R}^d\right)$ for some $n \geq 1$. Then the Taylor series expansion about $z \in \mathbb{R}^d$ is given by

$$u(z+b) = \sum_{|\beta|<n} \frac{a^{\beta}}{\beta!}\left(D^{\beta} u\right)(z) + (\mathcal{R}_n u)(z, b), \tag{A.2}$$

where $\mathcal{R}_n u$ is the integral remainder term

$$\begin{aligned} (\mathcal{R}_n u)(z, b) &= n \sum_{|\beta|=n} \frac{b^{\beta}}{\beta!} \int_0^1 s^{n-1}\left(D^{\beta} u\right)(z + (1-s) b)\, ds \\ &= n \sum_{|\beta|=n} \frac{b^{\beta}}{\beta!} \int_0^1 (1-s)^{n-1}\left(D^{\beta} u\right)(z + sb)\, ds, \end{aligned} \tag{A.3}$$

which satisfies the estimates

$$|(\mathcal{R}_n u)(z, b)| \leq \frac{d^{\frac{n}{2}}\, |b|^n}{n!} \max_{\substack{|\beta|=n \\ y \in [z, z+b]}} \left|\left(D^{\beta} u\right)(y)\right|, \tag{A.4}$$

and

$$|(\mathcal{R}_n u)(z, b)| \leq |b|^n \sum_{|\beta|=n} \max_{y \in [z, z+b]} \left|\left(D^{\beta} u\right)(y)\right|. \tag{A.5}$$

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover Publications, 1970.

[2] R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[3] J. Deny and J. L. Lions, *Les espaces du type de beppo levi*, Ann. Inst. Fourier **5** (1954), 305–370 (French).

[4] N. Dyn, *Interpolation and approximation by radial and related functions*, Approximation Theory VI (Chui C., Schumaker L., and Ward J., eds.), Academic Press, 1989, pp. 211–234.

[5] Stein E.M. and Weiss G., *Fourier analysis and euclidean spaces*, Princeton University Press, Princeton, 1971.

[6] I. M. Gelfand. and N. Ya. Vilenkin, *Generalised functions*, Academic Press, New York, 1964.

[7] R. L. Hardy, *Multiquadric equations of topography and other irregular surfaces*, J. Geophysical Res. **76** (1971), 1905–1915.

[8] ______, *Theory and applications of the multiquadric-biharmonic method*, Comput. Math. Appl. **19** (1990), 163–208.

[9] W. Light and H. Wayne, *Error estimates for approximation by radial basis functions.*, Approximation theory, wavelets and applications (Maratea, 1994), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 454, Kluwer Acad. Publ., Dordrecht, 1995, pp. 215–246.

[10] ______, *On power functions and error estimates for radial basis function interpolation*, J. Approx. Th. **92** (1998), no. 2, 245–266.

[11] ______, *Spaces of distributions, interpolation by translates of a basis function and error estimates*, Numer. Math. **81** (1999), no. 3, 415–450.

[12] W. Madych and S. Nelson, *Multivariate interpolation and conditionally positive definite functions*, Approx. Theory Appl. **4** (1988), 77–89.

[13] ______, *Multivariate interpolation and conditionally positive definite functions ii*, Math. Comp. **54** (1990), 211–230.

[14] P. Malliavin, H. Airault, L. Kay, and G. Letac, *Integration and probability*, Springer-Verlag, Berlin, 1995.

[15] C. Micchelli, *Interpolation of scattered data: Distance matrices and conditionally positive definite functions*, Const. Approx. **2** (1986), 11–22.

[16] B. Petersen, *Introduction to the fourier transform and pseudo-differential operators*, Pitman, 1983.

[17] M. Reed and B. Simon, *Methods of modern mathematical physics i: Functional analysis*, Academic Press, New York, 1972.

[18] R. Schaback, *Multivariate interpolation and approximation by translates of a basis function*, Approximation Theory VIII, Vol I; Approximation and interpolation (C.K. Chui and Larry Schumaker, eds.), World Scientific Publishing Co., 1995, pp. 491–514.

[19] I. J. Schoenberg, *Metric spaces and completely monotonic functions*, Ann. of Math. **39** (1938), 811–841.

[20] L. Schwartz, *Theorie des distributions*, Hermann, Paris, 1966.

[21] V. Vladimirov, *Equations of mathematical physics*, Marcel Dekker, New York, 1971.

[22] P. Y. Williams, *A weight function theory of zero order basis function interpolants and smoothers.*, Prepublication: http://au.arxiv.org/.

[23] K. Yosida, *Functional analysis*, Springer-Verlag, Berlin, 1958.

[24] ______, *Functional analysis*, Springer, Berlin-Heidelberg-New York, 1968.

[25] W. Zong-min and R. Schaback, *Local error estimates for radial basis function interpolation to scattered data*, J. Numer. Anal. **13** (1993), 13–27.